DOCUMENT RESUME

ED 240 313 CE 038 157

TITLE Assess Student Performance: Knowledge. Second Edition. Module D-2 of Category D--Instructional Evaluation. Professional Teacher Education Module Series.

INSTITUTION Ohio State Univ., Columbus. National Center for Research in Vocational Education.

SPONS AGENCY Department of Education, Washington, DC.

REPORT NO ISBN-0-89606-143-4

PUB DATE 84

NOTE 71p.; For related documents, see ED 228 460, ED 236 356, and CE 038 158.

AVAILABLE FROM American Association for Vocational Instructional Materials, 120 Driftmier Engineering Center, University of Georgia, Athens, GA 30602.

PUB TYPE Guides - Classroom Use - Materials (For Learner) (051)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS *Academic Achievement; Behavioral Objectives; *Cognitive Measurement; *Competency Based Teacher Education; Evaluation Criteria; *Evaluation Methods; Learning Modules; Measurement Techniques; Postsecondary Education; *Student Evaluation; Teaching Skills; *Test Construction; Test Items; Vocational Education Teachers

ABSTRACT

This module, one in a series of 127 performance-based teacher education learning packages focusing upon specific professional competencies of vocational teachers, deals with assessing student performance. Addressed in the individual learning experiences are the following topics: understanding the considerations involved in assessing student cognitive performance; constructing multiple-choice, matching, completion, true-and-false, essay, oral, case study, and problem-solving items to measure the achievement of selected student performance objectives; and assessing student cognitive performance in an actual teaching situation. Each learning experience contains an objective, instructional text, one or more learning activities, and a feedback activity. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED240313

MODULE
D-2

Assess Student Performance: Knowledge

Second Edition

Module D-2 of Category D—Instructional Evaluation
PROFESSIONAL TEACHER EDUCATION MODULE SERIES

The National Center for Research in Vocational Education
The Ohio State University

Key Program Staff:
James B. Hamilton, Program Director
Robert E. Norton, Associate Program Director
Glen E. Fardig, Specialist
Lois G. Harrington, Program Assistant
Karen M. Quinn, Program Assistant

Second Edition. Copyright © 1984 by The National Center for Research in Vocational Education, The Ohio State University, 1960 Kenny Road, Columbus, Ohio 43210.

Copyright is claimed until full term. Thereafter all portions of this work covered by this copyright will be in the public domain.

This work was developed under a contract with the Department of Education. However, the content does not necessarily reflect the position or policy of that Agency, and no official endorsement of these materials should be inferred.

1984

ISBN 0-89606-143-4

Published and distributed by the **American Association for Vocational Instructional Materials (AAVIM)**, 120 Driftmier Engineering Center, University of Georgia, Athens, Georgia 30602, (404) 542-2586.

CE038157

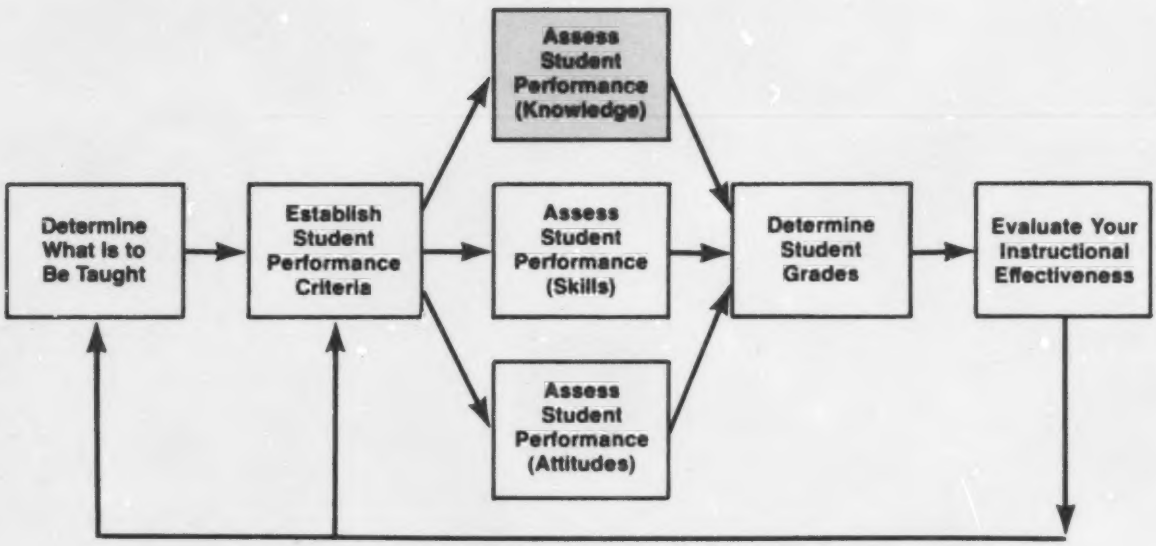
U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."


Determine What Is to Be Taught
Establish Student Performance Criteria
Assess Student Performance (Knowledge)
Assess Student Performance (Skills)
Assess Student Performance (Attitudes)
Determine Student Grades
Evaluate Your Instructional Effectiveness

INSTRUCTIONAL EVALUATION PROCESS

INTRODUCTION

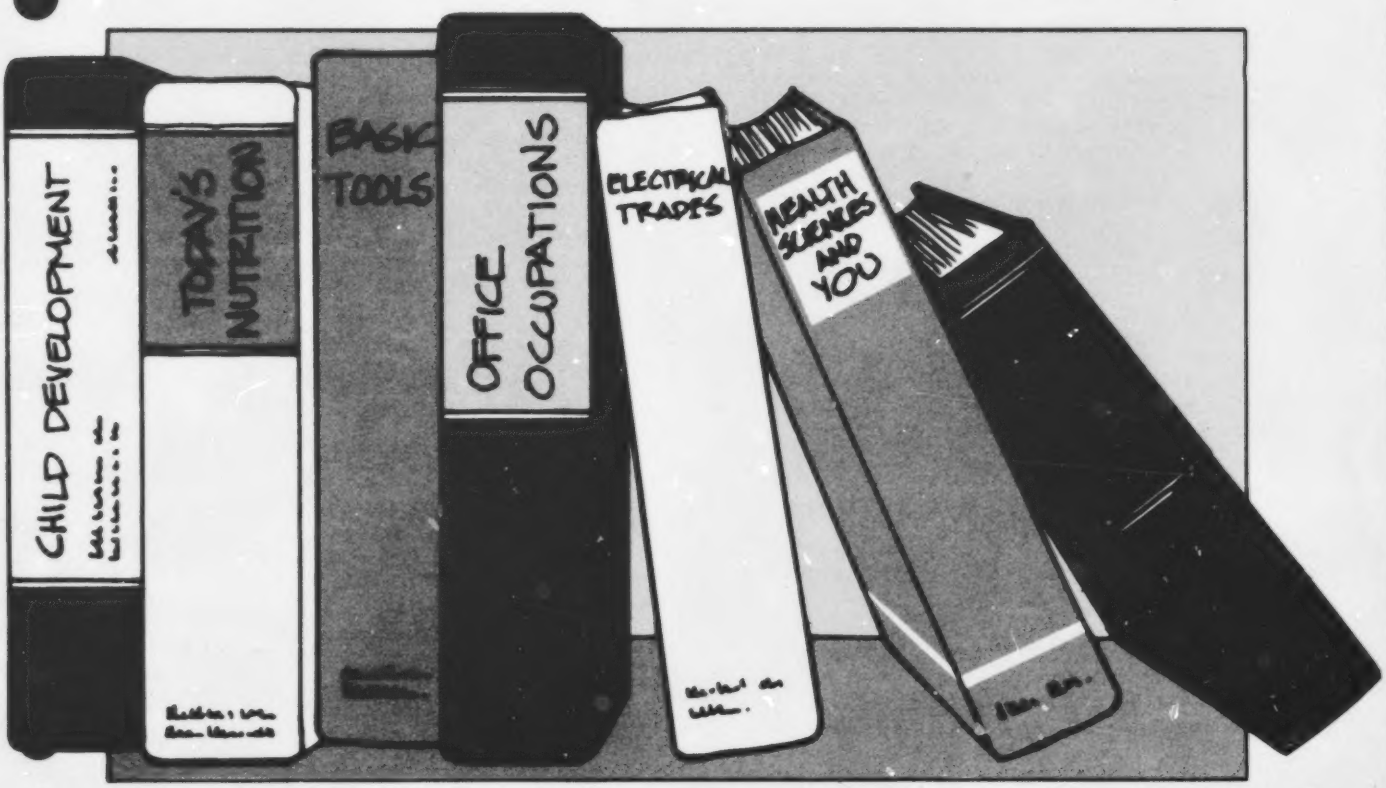
How will you know whether your students have acquired the technical knowledge implied or stated in the student performance objectives for the program? How can you be sure that they have learned the required theory, facts, data, or other information? The answer is that you can assess their knowledge level using some type of cognitive test items: true-false, multiple-choice, matching, completion, essay, or oral items.

Of course, measurement of student knowledge is only one part of the picture. In fact, there are three kinds of performance—knowledge, skills, and attitudes—that you will want to measure, and these are not truly discrete. In performing a skill, for example, a student is not only demonstrating psychomotor skill, but also certain attitudes (e.g., concern for safety and cooperation) and the possession of certain knowledge (e.g., the proper procedure to follow in making change for a customer).

However, there are many times when the measurement of knowledge alone is important. You may want to determine whether students **know** the steps in a procedure before allowing them to carry out that procedure for the first time. You may need to determine if students have understood a particular reading assignment or class lecture. You may want to test their ability to solve practice problems. And so on.

By assessing student knowledge regularly, you can keep them informed concerning their progress in learning the technical knowledge required for the occupation. You can determine their readiness to go on to subsequent learning activities. Assessment can also provide you with valuable information about the effectiveness of your own teaching—whether your students are learning, or whether you need to change your teaching strategies. And assessment also provides a basis for assigning grades in traditional vocational-technical programs.

This module is designed to help you gain the skills you need (1) to construct reliable, valid, and usable test items that will assess your students' achievement of the cognitive student performance objectives that make up your program; (2) to put those items together into an effective testing device; and (3) to create a testing environment that will allow students to demonstrate their knowledge.



ABOUT THIS MODULE

Objectives

Terminal Objective: In an actual teaching situation, assess student cognitive (knowledge) performance. Your performance will be assessed by your resource person, using the Teacher Performance Assessment Form, pp. 67–68 (*Learning Experience VIII*).

Enabling Objectives:

1. After completing the required reading, demonstrate knowledge of the considerations involved in assessing student cognitive performance (*Learning Experience I*).
2. After completing the required reading, construct five multiple-choice items to measure the achievement of selected cognitive student performance objectives (*Learning Experience II*).
3. After completing the required reading, construct a matching item to measure the achievement of selected cognitive student performance objectives (*Learning Experience III*).
4. After completing the required reading, construct five completion items to measure the achievement of selected cognitive student performance objectives (*Learning Experience IV*).
5. After completing the required reading, construct ten true-false items to measure the achievement of selected cognitive student performance objectives (*Learning Experience V*).
6. After completing the required reading, construct three essay items and two oral items to measure the achievement of selected cognitive student performance objectives (*Learning Experience VI*).
7. After completing the required reading, construct one case study or problem-solving item to measure the achievement of a selected cognitive student performance objective (*Learning Experience VII*).

Resources

A list of the outside resources that supplement those contained within the module follows. Check with your resource person (1) to determine the availability and the location of these resources, (2) to locate additional references in your occupational specialty, and (3) to get assistance in setting up activities with peers or observations of skilled teachers, if necessary. Your resource person may also be contacted if you have any difficulty with directions or in assessing your progress at any time.

Learning Experience I

Optional

Reference: Green, John A. *Teacher-Made Tests*. Second Edition. New York, NY: Harper & Row Publishers, 1975.

Reference: Gronlund, Norman E. *Measurement and Evaluation in Teaching*. Fourth Edition. New York, NY: Macmillan Publishing Co., 1981.

Reference: Karmel, Louis J., and Karmel, Marylin O. *Measurement and Evaluation in the Schools*. Second Edition. New York, NY: Macmillan Publishing Co., 1978.

Learning Experience II

Optional

Peers to critique your multiple-choice items.

Learning Experience III

Optional

Peers to critique your matching item.

Learning Experience IV

Optional

Peers to critique your completion items.

Learning Experience V

Optional

Peers to critique your true-false items.

Learning Experience VI

Optional

Peers to critique your essay and oral items.

Learning Experience VII

Optional

Peers to critique your case study or problem-solving item.

Learning Experience VIII

Required

An actual teaching situation in which you can assess student cognitive (knowledge) performance.

A resource person to assess your competency in assessing student cognitive (knowledge) performance.

General Information

For information about the general organization of each performance-based teacher education (PBTE) module, general procedures for its use, and terminology that is common to all the modules, see About Using the National Center's PBTE Modules on the inside back cover. For more in-depth information on how to use the modules in teacher/trainer education programs, you may wish to refer to three related documents:

The *Student Guide to Using Performance-Based Teacher Education Materials* is designed to help orient preservice and inservice teachers and occupational trainers to PBTE in general and to the PBTE materials.

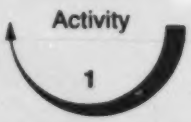
The *Resource Person Guide to Using Performance-Based Teacher Education Materials* can help prospective resource persons to guide and assist preservice and inservice teachers and occupational trainers in the development of professional teaching competencies through use of the PBTE modules. It also includes lists of all the module competencies, as well as a listing of the supplementary resources and the addresses where they can be obtained.

The *Guide to the Implementation of Performance-Based Teacher Education* is designed to help those who will administer the PBTE program. It contains answers to implementation questions, possible solutions to problems, and alternative courses of action.

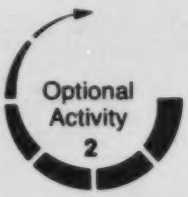
Learning Experience I

OVERVIEW

After completing the required reading, demonstrate knowledge of the considerations involved in assessing student cognitive performance.



You will be reading the information sheet, Student Cognitive Performance Assessment, pp. 6–16.



You may wish to read one or more of the following supplementary references: Green, *Teacher-Made Tests*, pp. 19–41; Gronlund, *Measurement and Evaluation in Teaching*, pp. 65–120; and/or Karmel and Karmel, *Measurement and Evaluation in the Schools*, pp. 397–409.

You will be demonstrating knowledge of the considerations involved in assessing student cognitive performance by completing the Self-Check, pp. 17–18.

You will be evaluating your competency by comparing your completed Self-Check with the Model Answers, pp. 19–20.

Activity 1

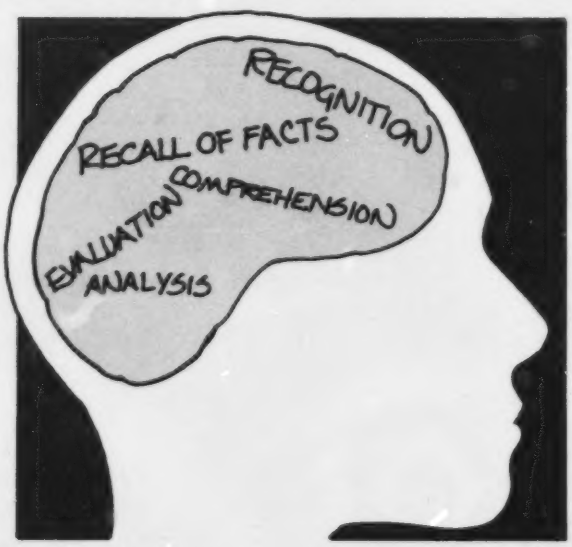
Assessing student cognitive performance, under one name or another, is as old as education. Experts agree that it is one of the vocational teacher's most important responsibilities. For information on why, when, and how to assess your vocational students' cognitive performance, read the following information sheet.

STUDENT COGNITIVE PERFORMANCE ASSESSMENT

Vocational-technical instruction should be based on identified student performance objectives, which state the intended outcomes of the educational process in terms of the specific knowledge, skills, and attitudes students need for entry into the world of work.[1] Learning in all three domains—knowledge (cognitive), skills (psychomotor), and attitudes (affective)—is important, but in this module we are concerned with the cognitive domain.

Benjamin Bloom[2] divides the cognitive domain into six levels. Cognitive objectives at the lowest level (i.e., knowledge) require students to recognize or recall correct facts, data, or information. Such objectives usually call for students to list, define, identify items, or in some other way recognize or recall a particular piece or body of information.

Student performance objectives at the higher levels in the cognitive domain require students to do more than simply recognize or recall factual information correctly. They also require students to **use** that information in some way. Objectives at the second level of the cognitive domain (i.e., comprehension), for example, may call for students to summarize, interpret, translate, or paraphrase facts, data, or information. Objectives at still higher levels may require application, analysis, synthesis, or evaluation of factual information.



Student cognitive performance assessment is the process of determining whether students have achieved these cognitive performance objectives. It is the process of testing students' knowledge and their ability to use that knowledge in a manner appropriate to your occupational area.

Purposes of Assessment

Information about students' cognitive performance—what knowledge they have acquired and how they can use that knowledge—can help in providing high-quality instruction for the students enrolled in your program. You and other people can put this information to several important uses.

Student readiness. You can use this information to determine whether your students are ready for a particular learning activity. Often, students must know one thing before they proceed to another activity. Let's say, for example, that a home economics teacher is demonstrating to students how to separate egg yolks from whites. The teacher has finished showing them the procedure and is now ready to have them try their skill on some real eggs.

Before the teacher turns students loose on those eggs, however, he or she might want to make sure that they're ready for the task. Eggs do cost money and they can make a nasty mess if mishandled. The teacher would be likely to quickly test students' knowledge of the procedure—their readiness for the hands-on task. He/she would want to ensure that students know (1) the **steps** in the procedure (e.g., first you break the egg); and (2) the **critical factors** in any steps (e.g., you must break the egg carefully, so that the two halves of the shell remain intact).

For that matter, you will undoubtedly test students' knowledge of safety rules and practices before students actually begin to use the tools and equipment in the laboratory. And, you will often test students' knowledge of theory before they proceed to practical exercises using that theory. Perhaps

1. To gain skill in developing student performance objectives, you may wish to refer to Module B-2, *Develop Student Performance Objectives*. To gain skill in determining the performance criteria that underlie your measurement of student achievement, you may wish to refer to Module D-1, *Establish Student Performance Criteria*.

2. *Benjamin S. Bloom, ed., Taxonomy of Educational Objectives, Handbook I: Cognitive Domain* (New York, NY: David McKay Company, 1956). Reprinted by permission of David McKay Company, Inc.

carpentry students are ready to begin construction projects including the purchase of lumber. The carpentry teacher would probably want to test students' knowledge of formulas for converting running feet to board feet before they make their purchases.

Instructional improvement. You can also use the information you get through assessing student cognitive performance to improve your instruction. Let's take the egg-separating example again. The teacher might test students' knowledge of the procedure and discover that they don't know what he/she thought they did. This could indicate that the instructional activities need improvement. Perhaps the demonstration was unclear or confusing. Perhaps some students were just sitting too far away to see what the teacher was doing. Finding out that the students haven't learned what you were teaching can alert you to possible weaknesses in your instruction.

Information about student progress. Other people can make good use of this information about students' learning, too. Students, for instance, need to know how they are progressing in your program. They will want to know if they are successfully acquiring the technical knowledge they will need on the job.

Providing students with feedback concerning their cognitive performance can help them identify their own weaknesses so that they can work more successfully to acquire needed technical knowledge. Positive feedback on the progress they are making can help to motivate them to progress further.

Administrators will also be interested in the information you gain by assessing student cognitive performance. One of your responsibilities, as a vocational-technical instructor, will no doubt be to report students' progress or achievement. Almost all vocational teachers must submit student grades or other progress reports at specified intervals. Part of students' overall progress, of course, is their progress in acquiring needed technical knowledge.

The information about student progress that you report to administrators may also be shared with other people in the school or community. Counselors might use this information in making placement decisions with students. Parents of secondary students are usually concerned with how their children are doing in the vocational-technical program.

Finally, employers frequently want to know about your students' competence. They commonly want to know students' course grades. They might even want to talk to you about the technical content of your program or about the knowledge, attitudes, and skills of a particular student. Of course, you can't tell potential employers what your students know, and what they can do with what they know, until you have found out yourself by assessing their cognitive performance.

Assessment Techniques

You might assess your students' cognitive performance at almost any time. It would not be exaggerating to say that you will assess cognitive performance, in one way or another, for one purpose or another, every day in your classes. Ensuring that students have mastered the technical content of your program should be an ongoing part of your instruction.

There are a number of different ways to assess student cognitive performance. In some situations, you might use very **informal** means to do the job. You might, for example, give a short oral quiz, requiring students to provide brief written answers that you can quickly check, during class, before they move into the laboratory. A final course examination, on the other hand, would almost certainly involve a full-blown, **formal** written test.

Regardless of how formal or informal the testing situation is, however, the test items that you will use to measure student knowledge will be of two types: objective and subjective.

Some kinds of cognitive test items are called **objective** by testing experts because scoring them is almost entirely an objective process. The following types of items are considered to be objective:

- Multiple-choice
- Matching
- Completion
- True-false

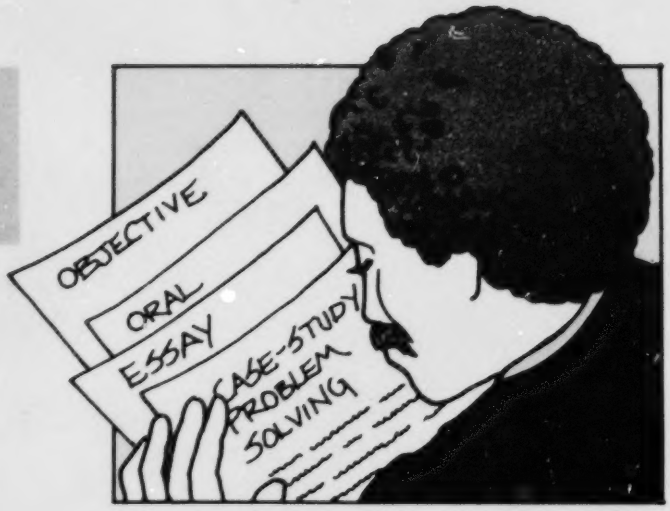
Scoring these items is primarily objective because the scorer (i.e., you) does not need to use judgment to determine whether a student's answer is right or wrong. Students' answers can be compared to a simple scoring or answer key. If the key says that the answer to an item should be "true," a student's answer either does or does not agree. Neither opinion nor judgment are needed for scoring.

Other kinds of test items are called **subjective** because they do require the use of judgment and interpretation in scoring answers. The following types of items are considered to be subjective:

- Essay
- Oral

As you are no doubt aware, scoring essay items is not as simple a matter as scoring objective items. Is a given answer the right answer? Is it partially right? If an essay question is worth 20 points, how many points is a partially right answer worth? Scorers must use their own judgment and opinion to resolve these questions in scoring subjective items.

There is another basic difference in test items. Some items require students to **supply** the answer from their own memory. The subjective types of items



are both supply tests, as you certainly know if you've ever taken an essay test. Of the objective types of test items, only one is of the supply type—completion items. Completion items require students to fill in the blank or give a short answer (word, numeral, symbol, or phrase) to a question, again from their own memory.

The other types of objective test items (multiple-choice, matching, and true-false), however, are **selection** types. In all of these, students are given two or more possible answers from which to select the correct answer or answers. Selection types test students' ability to recognize correct information, while supply types test students' ability to recall correct information.

Finally, a word about terminology. You may have noticed in the previous explanation that we talked about test **items** rather than test **questions** and about multiple-choice **items** rather than multiple-choice **tests.** The reasons that the word *item* is the preferred word in both cases are as follows:

- Items on a test may or may not be stated in interrogative—question—form. True-false items, for example, are not questions but declarative statements.
- A given test may include many different types of items. While you might devise a test that included, for example, only essay items (i.e., an essay test), you will also devise tests that include one section of essay items, one of true-false items, one of multiple-choice items, or some other combination.

Consequently, the word *item* is a more precise descriptor in both situations.

The Good Test: In Theory

What makes tests good or bad? What qualities should your tests have? Objective vs. subjective items, supply vs. selection items—how do you choose which to use? There is a wealth of literature in which testing experts agree that a good test should have the following characteristics.

The test must be valid. Validity is the extent to which a test measures what it is **supposed** to measure. To the extent that a test measures what it is supposed to, it is valid. This notion seems so simple and straightforward that you might wonder why it is even mentioned. How could you possibly go wrong?

You might go wrong if your test measures more than you intended. When you assess student cognitive performance, you want your tests to measure students' knowledge of **specific** facts, data, or information. Your test becomes less valid if it also measures other knowledge as well.

Assume, for example, that it is very important that students have read an assigned chapter before moving on to the next learning activity. You thus want to administer a test to ensure that they have read and understood the assigned material. One of your test items is the following multiple-choice item:

What country is currently pioneering work in the area of friction welding?

1. **India**
2. **Greenland**
3. **Nicaragua**
4. **Soviet Union**

Some students would be able to answer this item correctly using general knowledge, whether they had read the assigned material or not. Some students would be likely to know that, of the four countries listed, the Soviet Union has by far the most technologically advanced industrial economy. The Soviet Union would, therefore, be the most likely of the four countries to be engaged in pioneering work in a technological area. A guess based on this reasoning would be correct.

This item, then, measures students' general knowledge as well as their knowledge of the reading assignment. Could the item be repaired to measure what it is supposed to? If all the possible answers were equally plausible (i.e., all countries listed had technologically advanced industrial economies), general knowledge would be useless. Students would need to know the material covered in the assigned reading in order to answer the item. The item would then measure what it is supposed to, thereby increasing its validity.

Another way in which validity can be affected is if items on a test are answered by other items. For example, consider the following two completion items:

Harmful _____ live and grow on inanimate objects.

When harmful germs get inside a human body, _____ can begin.

The first correct answer—germs—is provided in the second item. Students who are discerning enough to recognize this can do well on the test without having truly learned the material. Thus, test validity is reduced.

The validity of a test can also be reduced if students must use skills other than those you intend to test in order to answer the items. For example, any written test requires students to use reading skills. Essay tests require the use of writing skills. Oral tests require the use of oral communication skills.

The need for students to use these other skills can be a serious problem, causing reduced validity, if some students cannot read or understand the items to begin with. If this happens, your test is measuring two things whether you intend it to or not—students' technical knowledge and their communication skills.

Students will need some level of communication skills to succeed in the world of work. We are not suggesting, by any means, that such skills are not important. But if you wish to measure **technical knowledge,** then you must ensure that your test allows students to show how much technical knowledge they actually possess. If a student's communication skills are weak, that problem should be identified and remediated, but it should not be allowed to **cloud** the measurement of technical skills.

The problem can be avoided if your items are carefully developed with students' communication skill levels in mind. In other words, written items should be at students' reading level. All items should be clearly and simply stated so that students' communication skills are sufficient to allow them to demonstrate their technical knowledge. In this way, the effects of communication skills on validity will not be significant.

The length of a test can also affect its validity. Let's say, for example, that you are preparing a final examination for your students covering the entire semester's technical content. You intend to test students' acquisition of a considerable body of technical knowledge and their ability to use that knowledge. Your final examination, then, must be long enough to cover all the knowledge and use of knowledge that you intend to test.

If your final examination has only five multiple-choice items, you are probably only testing students' knowledge of five small bits of information and not of the whole of the technical content covered during the semester. The five-item final examination thus would have very low validity. If the test is to measure knowledge of the many bits that comprise the whole technical content, it has to be long enough to include a representative sampling of all those bits.

On the other hand, a five-item test could have very high validity in a different situation. Perhaps you want to determine if students are ready to go on to hands-on activities after information activities. A five-item test might be quite sufficient if there are only five bits of information to check on. Since, in this case, you are sampling a small body of technical content, a short test should be sufficient.

The test should be reliable. Reliability is the **consistency** with which a test measures achievement. Experts in the theory of testing tell us that a valid test is always reliable—that is, it consistently measures what it is supposed to measure. Some theorists, in fact, treat reliability as a part of validity. In any case, it is important to note that a test can be reliable (i.e., can get a consistent measurement) even if it is not valid (i.e., does not measure what it is intended to measure).

This point can be illustrated with an extreme example. Let's say that you decide to measure students' knowledge of some information. For some unexplained reason—perhaps you've had a very hectic day—you write your test in German, even though none of your students understand a word of German. Now, obviously, this test would not be valid for any student who didn't speak German. Thus, it wouldn't be valid for your students. It would not measure what it is supposed to; instead, it would measure students' knowledge of the German language.

In spite of this, however, the test could be reliable. Let's say that you actually administer this German-language test to your students. There is a good chance that all of them would score zero. Furthermore, they would get this same score consistently. You could give them the same test again the next day and they would still score zero (unless they all spent the previous evening closeted with a German tutor). Hence, although this German-language test would have no validity for English speakers, it would yield reliable results—consistent scores.

One factor that has a great effect on test reliability is the subjectivity of the scoring. If a test has high reliability, you should be able to administer and re-administer it to a group of students and get roughly the same set of scores (making allowance for such

things as increased learning in the intervening time). Or, two different scorers should be able to sit down with a set of student tests, rate them independently, and arrive at very similar scores.

Subjective scoring, however, can lower the reliability of tests because it can allow inconsistent measurement. Two different teachers could each independently score student essay items and end up with completely different scores for the same answers. For that matter, one scorer could score a test one way one time yet score it completely differently another time.

Psychological studies have confirmed time and time again that inconsistent scoring can be a serious problem with subjective test items. Scorers must constantly make decisions—on the basis of their personal judgment, opinion, and preference—about the worth of students' answers.

Thus, it is important to objectify subjective test scoring insofar as possible. For example, **before** administering an essay test, you need to determine whether points will be deducted for incomplete sentences or for spelling errors. (Are these skills part of what you wish to measure? How many points will you deduct for each error or as a maximum?) You also need to prepare a scoring key, which lists the key points you expect to be made in each answer, and you need to determine how much value to assign to each key point covered. A thorough scoring key can minimize reliability problems with subjective tests.

Objective tests, on the other hand, do minimize or even eliminate subjectivity in scoring (hence, their name). If a true-false item is well written, for example, there can be only one answer to it. Further, no interpretation or judgment is required on the part of the scorer to determine if a student's answer is correct. The item is either true or false, and the student either did or didn't mark the correct response.

Different scorers, consequently, should have no difficulty in arriving at the same score for an individual student's test. Or, you could score the same set of papers twice and get exactly the same set of results. For that matter, with a scoring key, **anyone** should be able to score an objective test consistently, whether he/she knew one bit of the content tested or not.

As important as reliability is, however, it must be balanced with other concerns. Objective test items, in general, are good at measuring lower-level cognition (e.g., recall of facts). Higher-level cognition is more effectively measured through subjective test items. Thus, you must not, in the interests of reliability (and ease of scoring), use only objective measures. Rather, as mentioned before, you must ensure that your procedures for scoring subjective test items are as objective as possible.

The test should be usable. No matter how valid or reliable a test might be, it will be of little use to you if it is unreasonably difficult or time-consuming to prepare, administer, or score. A final examination with 1,000 objective items could be quite valid and reliable, for instance. Yet such a test would take so long to prepare, administer, and score, that it simply would not be practical to use.

Likewise, essay or oral items can in theory be used to test student recall of facts, data, or information—the lowest level of the cognitive domain. However, essay and oral items can be quite time-consuming. It takes students a longer time to write essay answers and you a longer time to score them. You can't give an oral item to the whole class at once (would they all answer together, in unison?), so considerable time would be needed for individual administration and scoring.

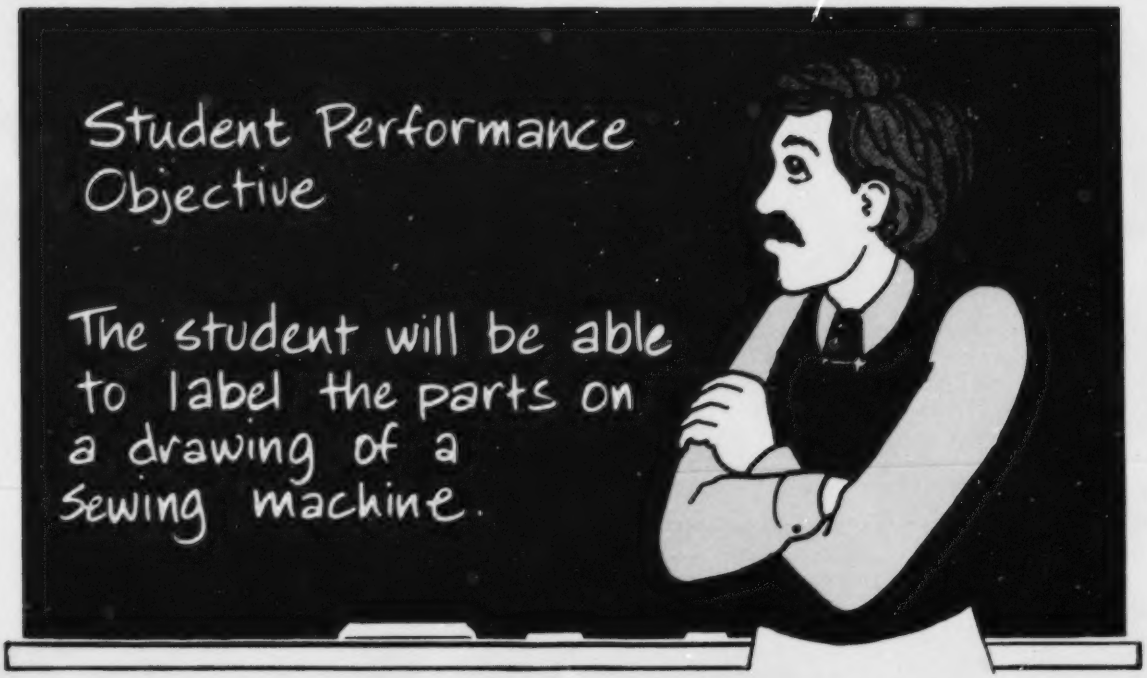
Consequently, essay and oral items would not be the most practical, usable means of assessing student cognitive performance at the recall level. Much more practical and usable in this case would be objective items.

Other factors may affect the usability of tests as well. Ultimately, you will need to use your own judgment to determine any other qualities that would make a test practical and usable in your own situation.

The Good Test: In Practice

How can you develop tests that are valid, reliable, and usable in your own program? The following guidelines can help you ensure that your tests measure what they are supposed to measure, and measure it consistently, while remaining practical and usable.

Base your test on student performance objectives. Just as your student performance objectives identify the knowledge, skills, and attitudes to be



taught, they also establish a basis for student evaluation. Logically, a given test should cover each of the student performance objectives that students should have achieved at a particular point. As you identify cognitive objectives to be covered in your test, you can also identify the specific criteria, described in the objectives, against which students' achievement should be measured.

You can then develop test items that also require students to demonstrate the required knowledge at the specified level (e.g., recall, synthesis). This will be of great importance in ensuring that your test has high validity—that it actually measures what it is supposed to.

The test should differentiate. If your test measures what it is supposed to measure, it should differentiate between students who know the material being tested and those who don't. If it does not differentiate, its validity will be low.

To tell if a test differentiates in this way, you will need to look at other indicators of student knowledge or achievement. These could be general in nature—the grades students usually get in your program, for example. If students who usually get good grades do well on a test, while students who usually get poor grades do badly on the test, it is probably reasonable to conclude that the test differentiates. Other general indicators of knowledge or achievement that you can use for comparison might be scores on standardized tests, general intelligence tests, or other tests of your own that you know to have high validity and reliability.

You should keep in mind, however, that many of these general indicators should be taken with a grain of salt. Standardized achievement tests and intelligence tests are often criticized by testing experts because of their cultural and language bias. Most appropriate for your purposes would be other specific indicators of knowledge of the particular content being tested.

For example, you might be able to compare students' test scores with their scores on daily quizzes covering the same content. Your own observation of student performance in the classroom or laboratory can often tell you which students know the material and which don't. For example, a student who does well on a test, and then also does well on hands-on practice activities that require using that specific knowledge, can reasonably be assumed to possess that knowledge.

Minimize the effects of communication skills. We have already discussed that every type of test item requires students to use some communication skills to understand and answer the items. A test becomes less valid if it requires students to use communication skills they don't have. When this happens, students cannot demonstrate their knowledge because they cannot communicate it to you. The test is not measuring what it is supposed to.

Thus, you should choose your test items carefully. For example, the essay item probably requires the highest level of communication skills. If a cognitive performance objective requires that students actually organize and present information themselves,

then essay items should, of course, be used. But if simple recall is required, you can use one of the objective types of items that does not require such high-level communication skills.

You can also minimize the effect of communication skills by writing clear, simple test items. Everyone is familiar with the trick question—the kind that you have to read and interpret in one particular way, and no other, to get the right answer. Questions like this, however, simply favor students who have greater skills in logic and in using the English language.

The more of these two skills the student has, the greater is the chance that he/she will successfully decode the trick item, divine your intention, and answer correctly. If your test is to measure what it is supposed to, your questions should be simple enough to be read in the same way by all students and answered correctly by all those who know the material.

This does not mean that you cannot test for fine distinctions in students' knowledge. On the contrary, you can and sometimes should do so. If your occupational area is one in which fine distinctions are important, you will certainly want to test students' knowledge of them. Furthermore, testing for these fine points can help your tests to differentiate between those who know the material and those who don't. Remember, nonetheless, that you can test for fine points in a clear, simple, direct manner.

The need to use communication skills in answering test items is a matter of particular importance for students with exceptional needs. A student with limited English proficiency, for example, might have considerable difficulty in understanding your test items or in writing answers to them. It is quite possible that such a student may know the information being tested perfectly well but be unable to communicate his/her knowledge to you.

Similarly, a visually impaired student might have difficulty reading a written test. A hearing-impaired student could have trouble with an oral test. And a physically impaired student might have problems in fitting written responses in the spaces provided. For these students, special modifications may be needed in the testing process.[3]

Provide clear, full, and simple directions. Difficult, complicated, incomplete directions can also reduce the validity and reliability of your test. If the directions are so difficult that some students cannot read them, validity will be reduced because of the effects of communication skills.

If the directions are unclear or incomplete, reliability will be reduced because the measurement won't be consistent. Mary Lou will answer an item this way because she thinks she is supposed to, while Johnny will answer it a completely different way, again because he thinks he's supposed to. Should this occur, students' scores will not truly indicate whether they knew the material or they didn't.

One good way to ensure that your test directions are clear, full, and simple is to test them out. Give the directions and a sample item to a student who is not going to be taking the test. Ask the student to read the directions and answer the item accordingly. You may want to test your directions with several different students, just to be on the safe side. Once you have developed and tested your directions for a particular kind of test item, you can reuse those directions for other items of the same kind.

In addition, it is often a good idea to include a sample item, with the correct answer appropriately marked, along with your test directions. An example provides students with another opportunity to be sure they understand your directions and answer appropriately. This is especially necessary if students are not used to the kind of test item, the way you want answers marked on their papers, and so on.

Your test directions should first explain administrative details pertaining to the whole test. You should tell students how much time is allowed for the test; how much time they should spend on various parts; and the point value of the test, the parts, and the individual items in the parts.

Students will also need to know whether they are to mark their answers on the test paper or on a separate answer sheet and exactly how to mark their answers. You should tell them whether they are to write the appropriate word in the blank in each item, circle the number of the correct response, place an *X* over the *T* in the column if the statement is true, or whatever. You should give separate directions for marking answers, with examples if appropriate, for each different kind of item that you use on the test.

Students should also be told what to do during the test if they should have additional questions. Their questions should be answered, but in such a way that it does not disturb the rest of the class. For example, you could ask students to raise their hands should they have questions, so that you can go to where they are sitting and answer such questions on an individual, and quiet, basis.

3. To gain skill in selecting and using testing techniques appropriate for students with exceptional needs, you may wish to refer to Module L-9, *Assess the Progress of Exceptional Students*.

Do not use too many different kinds of items. Authorities generally agree that you should use no more than three different kinds of items (e.g., true-false, essay, matching) on a single test. In addition, all items of the same kind should be grouped together.

Students need to develop a particular mindset in order to answer a particular kind of item. Achieving that mindset may be difficult and time-consuming for some students. You shouldn't, therefore, require students to change their mindset more often than is necessary—unless you also want to measure their ability to change modes of reasoning and answering rapidly on command.

If you find that you have prepared a test with too many kinds of items, you have two possible solutions, both of which are simple and easy. First, you might rewrite some of the items to convert them from one type to another. You could include these revised items with others of the same type, ending up with fewer kinds of items. Second, you could split your test into two tests and give students the two tests on two different days.

And note, please, that when you use different types of items on a single test, you should number all items consecutively (1, 2, 3, and so on) from beginning to end. Do not start numbering items in each new section with the number *1*. When you are discussing the items with students, either during or after the test, it is much simpler if there is only one item per number. Otherwise, there can be confusion—"Which number *1* do you mean, the essay item or the true-false item?"

The test should be just the right length. Determining the right length for a test involves a compromise between validity, on the one hand, and reliability and usability, on the other. We discussed previously how the validity of a test depends on how comprehensively the content is sampled. The larger the body of knowledge to be tested, the lengthier the test should be if it is to be valid.

In some situations, then, you might naturally be tempted to prepare very long tests. A final examination, for example, might test for knowledge of large quantities of information. You could logically conclude that the best approach is to use a large quantity of items. The more comprehensive the sample, the more valid the test—right?

That is true. However, if the test is too long, reliability can suffer. The longer the test, the more likely it is that students will become bored or fatigued. Fatigue and boredom are two student variables that affect reliability, because they affect the consistency of the students' performance in test-taking. If students are tired or bored, they will be less likely to remember what they know. They may confuse their memories of one bit of knowledge with another.

You will need to use your own judgment and your knowledge of your students, their needs, and their abilities to determine the right length for a given test. You should know your students well enough to know approximately at what point fatigue and boredom will begin to adversely affect their test-taking performance.

Remember, too, that the best compromise on length often depends on the situation. The test must be long enough to be valid for your purposes, yet not so long that reliability is reduced.

Watch for the guessing factor. There is always the possibility that a student could take a blind guess at any test item and get the right answer out of thin air. If a multiple-choice item has four choices, students have a 25 percent chance of getting the correct answer without even reading the item.

Students have a 50–50 chance on true-false items, which only offer two choices. Even in a matching item listing five inventions and the names of five inventors, the chances are only one in five—and sometimes the last item is free, by process of elimination.

One kind of guessing that you can and should prevent, however, is reasoned guessing. When you make a reasoned guess—an educated guess—you use whatever **clues** you can find in the item, together with your general knowledge, to figure out what the answer should be.

Multiple-choice items, for instance, often allow students to guess the right answer by process of elimination—"Well, I know for sure that it's not A, and I've never heard of B before, and C just doesn't sound right, so it must be D." To minimize the success of educated guesses, you will need to ensure that each of your items is carefully devised and free of clues to the correct response.

Now, in some situations, you may want to **ensure** that students cannot guess the correct answer because it is essential that they know the information being tested. Let's say that you're testing students' knowledge of safety rules and practices at the beginning of the year or term. You want to be absolutely sure that students **know** their safety rules and practices before they try using any tools, equipment, or machinery. In this situation, if they don't know, they might injure themselves.

The best approach when you need to be sure that they know is to avoid selection items. This means that you will have to decide whether to use essay items, oral items, or completion items. Completion items would be the most likely candidate. Your test can be valid, reliable, and usable—and students cannot guess.

Make good copies for everyone. If you are administering a written test, each student should have his or her own copy. So, be sure that you have enough well-produced copies of the test—neat, clean, and free from grammatical or typographical errors.

A sloppy production job (e.g., misspelled words, incomplete sentences, poor copies) invites students to approach the test haphazardly. It also creates one more opportunity for students to misunderstand your test directions or items.

Create a favorable testing environment. The physical environment in your classroom or laboratory should allow students to concentrate on taking the test. The room should be reasonably quiet and free from distractions. The temperature should be comfortable, and lighting should be adequate. Ventilation should be sufficient to keep the air in the room fresh.

The testing environment includes psychological factors as well. Some students become very anxious and apprehensive when taking a test, regardless of whether they know the material being tested or not. Or, students may be apprehensive because yours is a new kind of test for them—perhaps you're giving them their first essay items. You will often be able to minimize such apprehension by making the testing environment nonthreatening.

One way to do this is to make sure that students understand the purpose of the test. Is the test to be

graded, and are grades to be recorded? Or is the test an ungraded review of material? You should also tell students in advance exactly what material is to be covered in the test. Otherwise, some students will undoubtedly waste time trying to decide what to study—becoming, in the process, more and more apprehensive about the testing situation.

Another way to make the environment nonthreatening is to give students a practice test. This may be especially important if you are using a type of test item (e.g., essay) that your students are not used to. Practice tests give students the opportunity to gain experience with the type of test item and see that they can be successful at it.

Many testing experts also recommend putting easier questions at the beginning of your test. Their logic is that, since all students should be able to answer these easier questions, all students should be able to experience some initial success in taking the test, thereby decreasing their anxiety.

You will probably never be able to eliminate all unfavorable psychological factors in test-taking. You cannot control what might happen to your students outside your own classroom or laboratory. Any student may have just had an argument with friends or family and arrive to take the test in a distraught state. The possibility of these random psychological factors makes it all the more important that you do what you can to create a favorable testing environment.

Plan your test carefully. You will recall that the single most important characteristic of a good test is validity. A major factor affecting the validity of a test is how comprehensively it samples students' knowledge. Careful planning can help you to develop tests that are comprehensive samples of your students' knowledge.

The steps in planning a test are few and simple. First, you should identify the specific cognitive student performance objectives for which you want to measure student achievement. Second would be to list those objectives on a sheet of paper or planning form. This list of cognitive student performance objectives should describe the content to be covered by your test.

Third, you should review your daily lesson plans or learning packages, if necessary, to identify additional, more specific content to be included in the test. Sometimes, student performance objectives are not stated in enough detail to identify the specific content to be tested. Should you find that to be case, the greater amount of detail provided in your lesson plans or learning packages will be helpful.

Fourth, you will need to determine how many items to develop for each objective you have listed. Different objectives may well need to be covered by different numbers of items. How many items to include for a single objective depends on how important that objective is compared to the other objectives covered in the test.

Some indicators of this relative importance are (1) the amount of **time** spent on instruction for each objective, (2) the amount of **emphasis** placed on instruction for each objective, and (3) the amount of **material** covered by each objective.

There may be other indicators of relative importance to consider in your own situation. In any event, the more important or extensive an objective is, according to these indicators, the more items you should include to cover it. Furthermore, the relative importance of objectives should not come as a surprise to your students.

In determining the number of items to include for each objective, you are determining the relative weight to be given to different content in your test. If you choose, you can carry your planning forward another logical step. You can determine the point value or other weight to be assigned to each item or cluster of items.

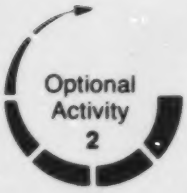
As you do so, it is very important to remember that point value should reflect the same relative importance of the objectives. If Objective A is covered by twice as many items as Objective B because it is twice as important, then the total point value of the items for A should be twice the point value of the items for B.

Sample 1 illustrates one planning form that you might find convenient to use in organizing your test planning efforts. You can list your cognitive student performance objectives or other descriptors of content in the left-hand column. The number of items to be developed for each objective can be listed in the second column; the type of items (e.g., multiple-choice), in the third column; and the total point value for these items, in the right-hand column.

SAMPLE 1

TEST PLANNING FORM

Cognitive Student Performance Objectives	No. of Items	Type of Items	Point Value



You may wish to read one or more of the following supplementary references:

- Green, *Teacher-Made Tests*, has many useful chapters. Chapter 2, pp. 19–41, discusses the levels of Bloom's taxonomy and the general principles of test construction.
- Gronlund, *Measurement and Evaluation in Teaching*, is a scholarly and detailed look at all aspects of evaluation. Chapter 3, pp. 65–92, examines validity in depth and discusses methods for determining validity. Chapter 4, pp. 93–120, treats reliability and formulas for determining coefficients of reliability.
- Karmel and Karmel, *Measurement and Evaluation in the Schools*, is a readable and informative compendium on evaluation. The chapter on teacher-made tests, pp. 379–409, compares the advantages and disadvantages of different test types.

The following items check your comprehension of the material in the information sheet, Student Cognitive Performance Assessment, pp. 6–16. Each of the four items requires a short, essay-type response. Please respond fully, but briefly, and make sure you respond to all parts of each item.

SELF-CHECK

1. In chatting with a colleague one day, the question of tests comes up. The colleague reports, with obvious pride, that no student ever scores 100 percent on a test in his program. How would you respond to this comment?

2. Another of your colleagues reports, in jest, that she plans to test her students' knowledge of a particular body of content soon. However, she has become tired of using that same old time-worn measurement tool, the written test. Instead, her tool will be a yardstick—she will take each student's height as a measurement of his/her knowledge.

   A silly suggestion, granted, but how **valid** would her technique be? How **reliable** would it be?

3. Today on your way out of the building, you overheard two students talking about a unit test they just took—a test consisting of one essay item. One student was pleased because she had concentrated her studies on the topic covered by the item.

   The other student was very angry and was complaining to the whole world about it. He had studied all the material covered in the unit and, consequently, didn't have as much to say about that one topic.

   Which student should get your sympathy—and why?

4. One day, a colleague stops to ask you a question. She is very concerned about the possible effects of communication skills on students' performance in taking written tests. Consequently, she wonders if it wouldn't be better to test knowledge **solely** through observation of students' performance in hands-on activities in the classroom or laboratory?

   "Wouldn't it be logical," she says, "to assume that, if students performed hands-on activities correctly, they must have known whatever factual information was required for the activity? Why not have students demonstrate their knowledge only in this way?" How would you respond to this suggestion?

Compare your written responses to the self-check items with the model answers given below. Your responses need not exactly duplicate the model responses; however, you should have covered the same **major** points.

MODEL ANSWERS

1. You could probably assume that there is something wrong with the teacher's tests. Even though you don't know anything about what type of test items the teacher uses or how he administers or scores his tests, the following line of reasoning should apply.

   Instruction in technical content should be appropriate for students' needs, abilities, and interests. Instruction should be of high quality. Tests should cover the same content as instruction. Test items should be clearly and simply stated. If these conditions have been met, then **some** students should score 100 percent at least **some** of the time.

   Not all students will get all answers correct, of course. Variations in individual needs, abilities, and interests will result in variations in test scores, as long as the test differentiates between those who know the material and those who don't. However, if no student ever gets 100 percent, you are forced to suspect that some of the three previous conditions are not being met.

   Possibly, none of the conditions are being met. If instruction is inappropriate or of poor quality, students' test performance will be just as poor. If students are being tested on content for which they received no instruction, they are not likely to do well. Finally, if test items are purposely ambiguous and misleading, then students will in all probability be misled by them and perform poorly on tests.

   You might give this teacher some recommendations for improving his tests. You could tell the teacher that he should base his tests carefully on his student performance objectives, just as he should base his instruction on them. The teacher should also avoid trick items that must be read and interpreted in only one obscure way. Attention to these two guidelines could improve the validity of his tests.

   Finally, the teacher should probably evaluate his instructional effectiveness. Although the problem might lie with his tests, it might just as well lie with his teaching.

2. How tall is knowledge? This teacher's test would have no validity. Whatever measurement tool the teacher uses must measure what it is supposed to in order to be valid. The tool she has chosen in jest could not measure knowledge. All the yardstick can measure is feet and inches, but knowledge doesn't come that way. How could Johnny be 5'6" smart?

   In spite of nonexistent validity, however, her technique would probably be highly reliable. The tool might not measure what it is **supposed** to, but it would probably measure what it does measure very **consistently.** She could give that same yardstick to another person and have him/her perform the same measurement on the same students. Both sets of results would most likely be very similar, if not exactly the same.

   Or, if she performed the same measurement on her students a second time, the results should again be very similar or the same. (This assumes that students haven't grown an inch or two smarter in the intervening time.) The problem of no validity would always remain, however. No matter how many times the teacher measured a student with the yardstick, she would still be measuring height and not knowledge.

3. In this instance, your sympathy should probably go to the student with the complaint. The young man most likely has a very good point. First of all, the instructor's directions in announcing the upcoming test should have made very clear the scope of material to be covered and the types of questions to be included.

   Assuming that the instructor simply indicated that a unit test was to be administered, then the young man would have been justified in studying, and expecting to be tested on, the full range of material covered in the unit.

   Any test must contain a comprehensive sample of the content being covered if it is to be valid. If the student's test was supposed to measure his knowledge of an entire unit, a sampling of all the key content of that unit should, in some reasonable way, be included in the test.

Frankly, it is difficult to imagine how all the content for an entire unit could be covered in a single test item. Admittedly, an essay item does have room for a good bit of content. But can **one** essay item require students to demonstrate all the learning outcomes described in all the cognitive student performance objectives included in that unit?

Logic alone should be sufficient to answer that question: No. The teacher in question has probably prepared a test with low validity because it was not a comprehensive sample of the content to be tested. Perhaps one essay item could measure achievement of one cognitive student performance objective—one day's lesson, for instance. But a whole unit's worth of content deserves a whole unit's worth of test.

4. Your answer to this colleague should be that the observation idea sounds good on the surface but could be extremely difficult to carry out in practice. The teacher's logic is pretty sound but probably just hasn't been carried far enough.

   You could construct a strong logical argument for using the observation approach. If it is certain that students must have certain knowledge to perform certain activities, then it is logical to conclude that successful performance indicates knowledge.

   Let's say that a business and office occupations student has successfully typed a given sequence of letters twenty times without error. This would surely seem to indicate that the student had **knowledge** of the location of the keys on the keyboard.

   But what if the student doesn't perform correctly? Should this be taken as an indication of no knowledge? Can specific bits of knowledge be linked definitively with specific aspects of performance? A great deal of time and work would need to be spent establishing the relationship between bits of knowledge and bits of performance.

   Furthermore, although ability to perform relies to a great extent on acquisition of the relevant knowledge base, not all vocational-technical content can be measured through performance. When the auto mechanics instructor provides information on the evolution of today's engine, he or she probably doesn't expect students to perform on those early engines described. But he/she might want to test their knowledge, through a cognitive test, of how today's engine evolved.

   In addition, instructors frequently want to test knowledge **before** permitting students to perform in order to further ensure that they perform competently and safely. Granted, a student can know the safety procedures and choose to ignore them in practice, so you would want to measure safety **performance** also. But first, you would want to make sure that **knowledge** of the procedures had been acquired.

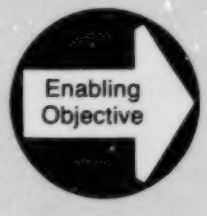
   Another objection to the colleague's approach is that it would have low usability compared to other forms of knowledge testing. Observation would require large amounts of instructor time, since each student's performance would need to be observed individually.

   In the final analysis, other types of tests are known to be effective in testing students' knowledge with far greater efficiency and practicality. Communication skills can indeed cause problems in testing—the teacher was rightfully concerned about this issue. Nonetheless, those problems can be successfully addressed by constructing test items and directions using a language level appropriate to the students taking the test.

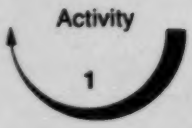
Level of Performance: Your written responses to the self-check items should have covered the same **major** points as the model answers. If you missed some points or have questions about any additional points you made, review the material in the information sheet, Student Cognitive Performance Assessment, pp. 6–16, or check with your resource person if necessary.

Learning Experience II

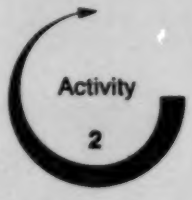
OVERVIEW



After completing the required reading, construct five multiple-choice items to measure the achievement of selected cognitive student performance objectives.

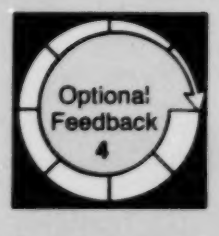


You will be reading the information sheet, Multiple-Choice Items, pp. 22–27.



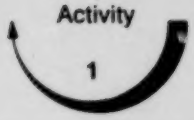
You will be selecting one or more student performance objectives that express or require cognitive learning.

You will be constructing five multiple-choice items to measure achievement of the student performance objectives you select.



You may wish to ask peers to give suggestions for the improvement of the multiple-choice items you develop.

You will be evaluating your competency in constructing multiple-choice items, using the Checklist for Multiple-Choice Items, p. 29.



Multiple-choice items are among the most popular and versatile objective test types. For information on how—and how **not**—to construct multiple-choice items, read the following information sheet.

MULTIPLE-CHOICE ITEMS

Multiple-choice items are selection items in which students are asked to choose the correct response to a given item from a short list of given responses. Each item consists of a **stem,** which supplies the central question or problem to which students are to respond. Each item also contains **responses,** from among which students must choose the correct response. Incorrect responses are called **distracters,** and the correct response is simply called the **answer.**

Multiple-choice items can appear in a variety of forms. First, the stem can be a complete direct question, as in the following:

Which of the following systems does the stomach belong to?

1. skeletal
2. respiratory
3. nervous
4. digestive

Or, the stem can be in the form of an incomplete statement, as follows:

A neutered male horse is called a

1. gelding
2. pinto
3. palomino
4. roan

Likewise, students can be asked to respond in a number of ways. Most commonly, students are asked to choose the **one** correct answer to the item. In this case, only one of the responses can be correct. All of the distracters must be completely incorrect. In the two examples used previously, there is only one correct answer.

Students can also be asked to identify the **best** response from among those listed. In this case, the distracters may be partially correct, but they are not as good as the answer. Students might be asked, as well, to identify **all** correct responses from among those listed.

Uses of Multiple-Choice Items

One reason that multiple-choice items are popular is that they are versatile. They can be used to measure student achievement in all six levels of the cognitive domain. It is true that the higher the level of knowledge to be measured, the more difficult it is to construct an item properly.

You will find, however, that with practice and care, you can write multiple-choice items that require students to go far beyond mere recognition of specific facts. Sample 2 contains multiple-choice items geared to each of the six levels of the cognitive domain.

Constructing Multiple-Choice Items

If you have ever tried to write multiple-choice items, you probably know how difficult that can be. Likewise, you probably know how frustrating it can be to try to figure out the meaning of an unclear multiple-choice item if you are the test-taker. You will find it helpful to use the following basic guidelines in constructing multiple-choice items.

The stem should be significant. Some testing experts say that the best stem for a multiple-choice item is one that would make a good completion item if you didn't supply responses to choose from. The first item in sample 2 has a stem that fits this rule. As you can see, the item could be converted to a completion item simply by dropping the choices and adding a blank at the end.

Not all experts agree that it is necessary to be quite this strict in writing stems, but all do agree that the stem should state the question or problem sufficiently so that only the right choice is justified as the answer. Only if the stem states the problem or question clearly can students be sure of what the item is asking and how to answer. The following is a **poor** example of a stem:

The ignition in a car

1. supplies current to the spark plugs
2. is part of the electrical system
3. consists of distributor, points, and coil
4. is likely to be electronic today

In fact, all the responses in this example are correct. Students would have an extremely hard time answering the item if they had to choose the **one** correct response. They would have to rely on their knowledge of the teacher's own way of thinking and teaching in order to guess which of these correct responses was the "right" one. This, of course, would

SAMPLE 2

MULTIPLE-CHOICE ITEMS FOR SIX COGNITIVE LEVELS

Level	Item
Knowledge	Frying is a form of cooking by contact with 1. dry heat 2. hot oil 3. flame 4. steam
Comprehension	A corner joint in which all crosscut surfaces are concealed is a 1. butt 2. dovetail 3. miter 4. rabbet
Application	If you have determined that you will need 10 six-foot lengths of 1″ × 6″ stock, how many board feet will you need to buy? 1. 15 2. 30 3. 60 4. 120
Analysis	The mare is to the stallion as the ewe is to the 1. ram 2. lamb 3. wether 4. mutton
Synthesis	If you were preparing a chocolate pudding using high heat, no stirring, and unbeaten eggs, the result would be 1. lumpy texture 2. smooth texture 3. curdling 4. soft consistency
Evaluation	Which of the following breakfast menus is nutritionally well balanced? 1. orange juice, frosted cereal, skim milk, apricot Danish 2. fried eggs, hash browns, donuts, coffee 3. tomato juice, coffee with cream, pancakes and syrup 4. orange juice, soft-cooked egg, whole wheat toast, skim milk

not test students' knowledge of ignitions, but rather their knowledge of the teacher's personality. In other words, the item would not be valid.

This stem could be improved by asking some questions about it. For example, what knowledge about the ignition are you trying to evaluate with the item—its function, its component parts, or its technology? The stem could be rewritten as follows to evaluate students' knowledge of the function of ignitions:

The function of the ignition in a car is to

If the stem contains opinions, say so. You may have occasion to test students' knowledge of opinion. You may justifiably want students in an automotive mechanics program to know that some experts think that diesels will replace internal combustion engines within 25 years. Consider the difference, however, between the following two examples:

Diesels will replace internal combustion engines within

a. 6 years
b. 10 years
c. 20 years
d. 25 years

According to the June 1982 issue of Diesel Weekly, *diesels will replace internal combustion engines within*

a. 6 years
b. 10 years
c. 20 years
d. 25 years

In the first example, students must essentially guess what your opinion would be so that they can agree with it. In the second, you have identified a specific authority as the source of the opinion, so that students can use their knowledge of this authority's opinion as the basis for their answer.

Don't end your stem with a give-away. Sometimes, the last word in the stem can give the answer away if you are not careful in writing your responses. In the following example, notice that the stem ends with the word *an*.

A physician who specializes in the structure, functions, and diseases of the eye is an

1. hematologist
2. optician
3. optometrist
4. ophthamologist

Many students will be able to rule out the first response just on the basis of their knowledge of the English language—you don't say "an hematologist." One way to remedy this error is to end the stem at the word *is* and to include the word *a* or *an*, as appropriate, in each response.

All responses for a single stem should use the same grammatical form. This rule is closely related to the preceding one. If your stem is an incomplete statement, each response should be stated in a form that correctly completes the statement in the stem. In the following example, students could again use their knowledge of the English language to rule out the second response.

Farmers rotate their crops in order to

1. spread out the work load
2. ease of marketing
3. conserve the soil
4. balance their diet

Once again, the stem does not combine with the second response to form a grammatically correct sentence in English.

All responses should be about the same length. Item analysis has repeatedly shown that the correct response tends to be longer than the distracters. Often, the correct answer must be qualified with extra words and phrases just in order to make it true. If your answers are usually longer than your distracters, however, test-wise students are provided with a clue to the correct response.

All distracters should be plausible. You might, for example, be writing a multiple-choice item to test students' knowledge of the name of an inventor (e.g., *The sewing machine was invented by*). If so, Christopher Columbus is not a plausible answer, at least for secondary or postsecondary students. You could expect any teenager or adult to be able to rule out Columbus on the basis of general knowledge. You would be presenting, again, a clue to the correct response. An excellent source of plausible distracters is the incorrect answers students have given to related completion items in the past.

Avoid using negative statements in the stem. Logically, it might well seem that there should be no problem in the following example:

Which of the following woods is not a hardwood?

1. maple
2. oak
3. cherry
4. pine

The question does not appear difficult—four woods are listed, one of them is not a hardwood, so which one is it? Research shows consistently, however, that students do less well on multiple-choice items that have a negative statement in the stem. Apparently, the small word *not* is often simply overlooked in the pressure of the testing situation.

Furthermore, most experts agree that it is less than ideal to emphasize negative learning. Why not, they say, rewrite the previous example so that the stem asks, *Which of the following woods is a softwood?* Now, the item emphasizes the positive fact that pine belongs to the softwoods.

The same experts also agree, however, that it may sometimes be necessary to test for negative knowledge (e.g., *Which of the following things should you NOT do when using a radial arm saw?*). When this is so, you should at least underline the negative word or write it in all capital letters, to make it stand out as much as possible.

Finally, in avoiding negatives, you should **never** write a stem containing a double negative, especially a combination of *not* and a negative adjective (e.g., *not impossible* or *not unlikely*). This, again, only serves to introduce irrelevant difficulty into the item. Such an item may test students' skill in using the English language more than their knowledge of technical content.

Avoid any pattern of response. Some teachers tend to put the correct response in the middle of the list, apparently because they feel that putting it in the first or last position will make the correctness of the answer too obvious. Test-wise students, however, will notice this pattern and use it to rule out incorrect responses.

Likewise, students are likely to notice any other pattern you might use. It is best to go back to each item after you have completed writing all items and to distribute all responses randomly. In that way, you can ensure that there will be no pattern of responses for students to use as a clue.

Don't give clues to one item in another. An obvious example of this would be the following two stems:

> ***The catalytic converter, in common use since 1975, was invented by***
>
> ***When did the catalytic converter first come into common use?***

The example probably seems extreme, especially since the two items are right next to each other. However, you can very easily write two such items and not notice it yourself when you have written several other items in between. It is best to review all items together to ensure that none of them contain clues to any of the others.

Avoid *all of the above* and *none of the above* responses. Using these responses can, once more, provide clues to experienced test-takers. Students can assume that, if they find two responses that they know to be correct, the answer has to be *all of the above*. On the other hand, if students can find one answer that they know is correct, they can automatically rule out *none of the above*. They may, therefore, select the correct response not through knowledge, but through the process of elimination.

Fit the item to the objective it covers. If your objective is for students to apply knowledge, then your multiple-choice item should actually test their ability to do that. The item must present a **new** situation in which students are to apply knowledge. If you use the same example on the test that you used in a classroom discussion, all you are testing is students' recall of your earlier example. You are not testing students' ability to use information in a situation different from the original learning context.

Use four or five responses in each item. The more responses there are in an item, the less likely it is that students will be able to get the correct answer by guessing or by process of elimination. Four or five items seems to be the best number. The element of pure chance is reduced, yet all the responses can still be read fairly quickly.

Keep all multiple-choice items together. If your test contains more than one kind of item, you should keep all the items of one kind together. This keeps to a minimum the number of times students have to get the right mindset for the type of item. They don't have to switch back and forth from a true-false item to a multiple-choice item to a completion item and back to another multiple-choice item. If you mix your types of items all together this way, you are also testing students' ability to switch rapidly from one mode of reasoning and answering to another.

Give clear, simple, complete directions. Your students are very likely to know that there are different ways to answer multiple-choice items. You might want them to choose the one correct answer, the one best answer of several that are all correct to varying degrees, all possible correct answers no matter how many, and so on. Consequently, it is important that you state clearly and completely how students are to respond.

You should also be sure to tell students how to mark their answers. Depending on how you plan to score the finished tests, you may want students to mark their answer sheet in different ways. For example, if tests are to be computer-scored, students will need to be told to use the computer answer form and a #2 pencil. Or you can ask students to place the correct letter or number in a particular blank, to mark an *X* through each correct response, or to circle the correct response, and so on.

Sample 3 presents examples of clear and simple directions for different kinds of multiple-choice items.

You will also need to use your own judgment and knowledge of your students to determine whether they should be given written or oral directions. If there is a chance that students will not understand written directions, you should give directions orally, by all means. If students are not accustomed to multiple-choice items or to taking your tests, it might be wise to provide both written and oral directions. You may even want to show students an example of how to answer.

Prepare a scoring key. Scoring students' responses will be much easier and more reliable if you use a scoring key. If, for example, you ask students to simply mark an *X* through the number or letter of each correct response, then your key can be a stencil-type. A stencil is a copy of the test with a hole punched through each correct response. When you place your key over each completed student test, each response showing through a hole should be marked with an *X* or the response is wrong.

A stencil may not work, however, if you ask students to circle each correct response. In that case, their circles might be bigger than the holes in the stencil, which would cause you to mark a correct response as incorrect.

A stencil is easiest to use when students use an answer sheet with columns of numbers or letters for marking answers. For this reason, some teachers always put such columns in one margin or the other. Students mark their answers in this column, and not on the choices actually listed in the item, as follows:

1 2 3 4 ***Muzzle velocity is ordinarily expressed in***

1. **yards per second**
2. **yards per minute**
3. **feet per minute**
4. **feet per second**

SAMPLE 3

DIRECTIONS FOR MULTIPLE-CHOICE ITEMS

Selecting a single correct response:

Directions: Each of the items below is followed by four possible responses. For each item, only one of the responses is correct; the others are incorrect. Select the one **correct** response for each item. Indicate your answer by placing an *X* over the number of the correct response.

EXAMPLE: The commonly accepted industrywide standard for beginning-level typists is

1. 40 words per minute, with a minimum of five errors
2. 40 words per minute, error-free
3. X 45 words per minute, with a maximum of three errors
4. 45 words per minute, error-free

Selecting a single best response:

Directions: Each of the items below is followed by four possible responses. For each item, any or all of the responses could be correct. Select the response that is **best** for each item. Indicate your answer by placing an *X* over the number of the best response.

EXAMPLE: You have just typed a letter for your supervisor, Mrs. Staunton. She gives it back to you and says that she has decided that she doesn't want to include the last paragraph in the letter after all. What would you do to remove that last paragraph:

1. use correction tape
2. cut and paste
3. X retype the letter
4. use correction fluid

Selecting all correct responses:

Directions: Each of the items below is followed by four possible responses. For each item, any or all of the responses could be correct. Select **all** the correct responses for each item. Indicate your answer by placing an *X* over the number of each correct response for each item.

EXAMPLE: You can thicken gravy with

1. X cornstarch
2. cornmeal
3. X wheat flour
4. baking soda

Select one or more student performance objectives in your occupational specialty that express or require cognitive learning and that lend themselves to the use of multiple-choice items to measure their achievement. (If you need help in identifying student performance objectives that require or express cognitive performance, check with your resource person.)

Construct five multiple-choice items to measure achievement of the student performance objectives you selected. Number each item for easy reference during feedback. Include directions and a scoring key.

You may wish to ask one or more of your peers who have taken or are taking this module to critique the multiple-choice items you have constructed. Discuss any suggested changes and then make any necessary revisions.

After you have constructed your multiple-choice items, use the Checklist for Multiple-Choice Items, p. 29, to evaluate your work.

CHECKLIST FOR MULTIPLE-CHOICE ITEMS

Directions: Place an X in the NO, PARTIAL, or FULL box to indicate that each of the following performance components was not accomplished, partially accomplished, or fully accomplished. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

Name _____

Date _____

Resource Person _____

LEVEL OF PERFORMANCE

The multiple-choice items meet the following criteria:	N/A	No	Partial	Full
1. A significant question or problem is stated in the stem of each item	☐	☐	☐	☐
2. All responses for each item are grammatically correct	☐	☐	☐	☐
3. All the response choices within each item are approximately the same length	☐	☐	☐	☐
4. An authority is quoted when the item contains opinion	☐	☐	☐	☐
5. There are no ambiguous statements in the stems or responses	☐	☐	☐	☐
6. All distracters for each item are plausible	☐	☐	☐	☐
7. The items are written at the language level of the students	☐	☐	☐	☐
8. The items do not contain clues to the correct responses for other items	☐	☐	☐	☐
9. The items are constructed to measure the level of knowledge specified in the objectives	☐	☐	☐	☐
10. Stems do not contain negative items	☐	☐	☐	☐
11. Correct responses are randomly distributed	☐	☐	☐	☐
12. Each item contains four or five responses	☐	☐	☐	☐
13. Clear, sufficient directions are provided	☐	☐	☐	☐
14. An appropriate scoring key is provided	☐	☐	☐	☐

Level of Performance: All items must receive FULL or N/A responses. If any item receives a NO or PARTIAL response, review the materials in the information sheet, Multiple-Choice Items, pp. 22–27, revise your items accordingly, or check with your resource person if necessary.

NOTES

Learning Experience III

OVERVIEW

After completing the required reading, construct a matching item to measure the achievement of selected cognitive student performance objectives.

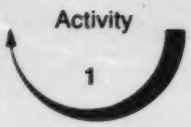
You will be reading the information sheet, Matching Items, pp. 32–34.

You will be selecting one or more student performance objectives that require students to recall or recognize information.

You will be constructing a matching item to measure achievement of the student performance objectives you select.

You may wish to ask peers to give suggestions for the improvement of the matching item you develop.

You will be evaluating your competency in constructing a matching item, using the Checklist for Matching Item, p. 37.



Another familiar and popular form of objective testing is the matching item. For information on what to do and what to avoid in constructing matching items, read the following information sheet.

MATCHING ITEMS

Like multiple-choice items, matching items are a form of selection item. The matching item consists of two lists of elements. One list (usually the one on the left) is made up of **premises,** while the other contains **responses.** The students' task is to match each premise with the appropriate response on the basis of a stated association or **relationship** between the two elements.

Uses of Matching Items

There are many kinds of associations or relationships that can be used as the basis for a matching item. Students could be asked to match inventions with their inventors, tools or pieces of equipment with their uses, terms with their definitions, principles with examples of their application, and so on. Since you will probably want to test students' knowledge of such associations or relationships in your own occupational area, you will find matching items to be useful.

Furthermore, matching items can be constructed using real objects or materials, pictures, drawings, or models. Instead of using words as your premises, you could use pictures for students to match responses with. You could, for example, give horticulture students pictures of different leaves to match with the names of the trees they come from.

Or, you could use one large, clear drawing or diagram of a radial arm saw, with arrows drawn to indicate different parts of the saw. Students would then match each part indicated by an arrow with the correct name from the list of responses.

You can also construct matching items with three parts instead of two. You could use a three-part matching item for terminology in health care occupations. The list of premises could consist of Latin or Greek terminology. The first list of responses could be literal translations of the premises, and the second list of responses could be the actual definitions of the terms. It is not recommended that you attempt to use matching items with more than three parts (one premise and two responses).

Matching items are really only appropriate, however, for measuring student achievement at the **lowest** cognitive level—knowledge. Experts in testing generally agree that matching items are not a valid measure of any of the higher cognitive levels.

Constructing Matching Items

While matching items may be familiar and popular, it is very important to take care in constructing them. When constructed carefully, matching items can be valid, reliable, and useful in evaluting student knowledge. The following are guidelines for writing good matching items.

Use a reasonable number of both premises and responses. Neither of these two lists should be either too long or too short. If there are too few elements, it becomes easier for students to guess correct answers—especially the last one—by process of elimination. (If there are only three items in each list, you can get them all right even if you only know two.) On the other hand, students have to spend too much time reading the list over and over if there are too many elements in the list.

Unfortunately, there is no single answer concerning how long the lists should be. Almost all the experts agree that five is the fewest number of elements you should put in a list. The maximum number recommended varies from ten to fifteen. It would appear safest to keep to the middle ground and have approximately ten elements in each list—certainly no fewer than five and no more than fifteen.

All premises or responses in one list should be homogeneous. Stated simply, this means that all the elements in a list, whether premises or responses, should be the same kind of thing. You might, for example, want students to match two different sets of things: technical terms with their definitions, and tools with their uses. This would best be done by constructing two different matching items. In one, students could match terms with definitions. In the other, they could match tools with uses.

The whole matching item should be on a single page. Students should not have to flip back and forth from one page to the next to scan all the elements in either list. This would probably confuse students and take too much time. If necessary, you should leave the bottom of one page blank and start your matching item at the top of a fresh page in order to fit it all on one page.

List enough responses so that some are left over. If you have exactly the same number of premises and responses, and each response is used only once, many students will get the last one correct automatically, by process of elimination. This reduces the validity of the item, of course.

You can avoid this problem in one of three ways. The first is to use more responses than premises—for example, six premises and nine responses. Then students still have four possible responses left when they get to the last premise. Listing two or three more responses than premises is sufficient.

The second way to avoid the problem is to use a single response as the correct answer for more than one premise. You might, for example, have ten premises and ten responses, with one response being the correct answer for two premises and one response not used at all. If you do so, you must tell students that a response can be used once, more than once, or not at all.

One final variation is possible. You can state in your directions that a response can be used once, more than once, or not at all. However, you can then list the same number of responses as premises, with each response being used once.

This is a legitimate variation, since it does fit your directions for using responses. It can be effective in eliminating guessing by process of elimination under one condition—that your matching tests **do** use a single response more than once often enough that students know that this is a real possibility. If you simply recite the formula for every matching item but never really use responses that way, students will soon learn that you don't really mean it.

List your responses in some logical order. If responses are listed in a logical order, students can find the response they are looking for more easily and quickly. If the responses are dates, they could be listed in chronological order. If they are names or terms, they could be in alphabetical order. Other possible logical orders might be increasing or decreasing size or importance.

The premises can be long, but the responses should be short. When students are actually taking the test, they usually read one premise and then scan the list of responses looking for the correct one. As a result, they end up reading and rereading the list of responses. To save time, the responses that they will be reading over and over should be short.

Avoid matching patterns. It is surprisingly easy to place your responses so that they form a pattern in relation to your premises. You might do this without realizing it. For example, you might automatically put a response in the bottom half of the list when its premise is in the top half. Or, if the premise is toward the middle you might tend to put the appropriate response in the middle also. Other patterns are also possible, and all should be avoided. Otherwise, some students will notice your patterns and find clues to correct responses.

One way to avoid matching patterns is to review the entire matching item after you have finished writing it. Look for patterns in the way elements match. If you find such patterns, you should use a procedure to randomize the order of the responses. You might roll dice, pick numbered slips of paper out of a hat, number the sides of a pencil and roll it, or perhaps use the randomization function of a calculator. Listing your responses in some logical order, as described earlier, may also serve to randomize them.

Provide clear, simple, and complete directions. You should not assume that your students will understand what to do with a matching item as soon as they see it. Nor should you assume that students will automatically know what the association or relationship between the premises and responses is. Students should also be told how to mark their answers.

Thus, in your directions, you should tell students what the relationship in the item is—for example, that they are to match breeds of swine with their identifying characteristics—and exactly how they should mark their answers. Sample 4 presents a matching item with directions for its use.

Prepare a scoring key. A scoring key always makes the job of scoring easier and helps to eliminate errors in scoring. You can just write the numbers and letters of the premises and the correct responses down the side of a piece of paper. Then, for quick and accurate scoring, you can lay this key next to the blanks in which students mark their reponses.

SAMPLE 4

MATCHING ITEM AND DIRECTIONS

Matching Test:
The Typewriter Keyboard

Directions: The list on the left describes the functions of different keys on the typewriter keyboard. The list on the right contains the names of different keys. Match the names of the keys with their correct functions by writing the letter of the correct response in the blank to the left of each function. Use each response only once; the first item is completed as an example.

	Functions	**Keys**
b 0.	Removes a letter, number, or symbol already printed	a. Back space
_____ 1.	Moves the typing element quickly across the keyboard to a position of your choosing	b. Correction key
_____ 2.	Moves the typing element one space to the left without printing anything	c. Margin release
_____ 3.	Prints capital letters (A) instead of lower-cased letters (a)	d. Return
_____ 4.	Moves the typing element one space to the right without printing anything	e. Shift
_____ 5.	Moves the typing element to the left margin and rolls the paper up to the next line	f. Shift lock
_____ 6.	Prints whole words or sentences in capital letters	g. Space bar
		h. Tab

Select one or more cognitive student performance objectives in your occupational specialty that require students to recall or recognize information.

Construct a matching item to measure achievement of the student performance objectives you selected. Include directions and a scoring key.

You may wish to ask one or more of your peers who have taken or are taking this module to critique the matching item you have constructed. Discuss any suggested changes and then make any necessary revisions.

After you have constructed your matching item, use the Checklist for Matching Item, p. 37, to evaluate your work.

NOTES

CHECKLIST FOR MATCHING ITEM

Directions: Place an X in the NO, PARTIAL, or FULL box to indicate that each of the following performance components was not accomplished, partially accomplished, or fully accomplished. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

Name _____

Date _____

Resource Person _____

The matching item meets the following criteria:	LEVEL OF PERFORMANCE			
	N/A	No	Partial	Full
1. Only homogeneous premises and responses are grouped within the item	☐	☐	☐	☐
2. There are not less than 5 and not more than 15 responses in the item	☐	☐	☐	☐
3. Premises are arranged for maximum clarity and convenience for the student	☐	☐	☐	☐
4. The responses are logically (e.g., alphabetically, chronologically) arranged	☐	☐	☐	☐
5. The directions clearly indicate the basis for matching and the method for marking answers	☐	☐	☐	☐
6. Real materials, pictures, drawings, or models are used when practical	☐	☐	☐	☐
7. Perfect one-to-one matching in the arrangement of premises and responses is avoided	☐	☐	☐	☐
8. All the premises and responses are on one page	☐	☐	☐	☐
9. There are more responses than premises, or single responses can be matched with more than one premise	☐	☐	☐	☐
10. Correct answers are randomly distributed	☐	☐	☐	☐
11. Responses are shorter than premises	☐	☐	☐	☐
12. An appropriate scoring key is provided	☐	☐	☐	☐

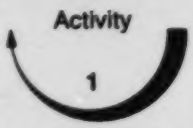
Level of Performance: All items must receive FULL or N/A responses. If any item receives a NO or PARTIAL response, review the material in the information sheet, Matching Items, pp. 32–34, revise your item accordingly, or check with your resource person if necessary.

NOTES

Learning Experience IV

OVERVIEW

After completing the required reading, construct five completion items to measure the achievement of selected cognitive student performance objectives.



You will be reading the information sheet, Completion Items, pp. 40–41.

You will be selecting one or more student performance objectives that require students to recall information.

You will be constructing five completion items to measure achievement of the student performance objectives you select.

You may wish to ask peers to give suggestions for the improvement of the completion items you develop.

You will be evaluating your competency in constructing completion items, using the Checklist for Completion Items, p. 43.

Completion items can be used to great advantage in evaluating student knowledge. For information on how to construct high-quality completion items, read the following information sheet.

COMPLETION ITEMS

Completion items are **supply** items—that is, students must supply their own answers instead of choosing from given possible answers. Completion items can appear in two forms. One form is simply a sentence with a blank in place of one or more missing words, as in the following example:

The unit used to measure electrical resistance is called the _____.

Another form consists of a complete question requiring a short answer (hence the name *short-answer test* sometimes used for a test using these items). Students write their answer in a blank provided at the end of the question, as in the following example:

What are the parallel beams that support the roof of a structure called?_____

Uses of Completion Items

Completion items test students' **recall** of knowledge. Hence, they are different from selection items (e.g., multiple-choice and matching) in which students need only to be able to **recognize** the correct answer. In this respect, completion items are similar to essay or oral items because all of these forms require students to supply answers from their own memory. The possibility of guessing thus is minimized.

Completion items can be used to test student recall of many different kinds of factual information. They can be used to evaluate students' knowledge of specifics, terminology, classifications, methodology, universals, principles, and so on. Furthermore, they can be used in any content area. However, completion items are most appropriate for use at the first—lowest—level of the cognitive domain.

Constructing Completion Items

Care should be taken in constructing completion items. If an item is not clearly stated, students may misinterpret it and answer incorrectly, even though they might know the correct answer. On the other hand, a poorly written item may contain clues that a test-wise student can use in determining the correct answer. Both situations should be avoided.

The following are suggestions for constructing completion items.

Use your own words. Many authorities feel that the main weakness of the completion item is that it only measures recall of rote learning, thus encouraging lower-level cognitive activity at the expense of higher-level activity. Lifting textbook quotes verbatim to use as completion items encourages rote learning even more.

Test only for significant bits of knowledge. You will need to use your own judgment to determine what knowledge is significant in your own area. Using that judgment, you can ensure that your completion items focus on significant knowledge in the area. Students might truly need to know that a particular technological development occurred in your field in 1962. It is unlikely, however, that students need to know that this development took place on a Wednesday morning.

Have each item clearly call for one—and only one—answer. Perhaps the most common mistake teachers make in writing completion items is in not being clear enough. Consider the following example:

The cathode ray tube with fluorescent screen was first introduced by K.F. Braun in _____.

How many possible answers are there to this item? The tube in question was first introduced by Braun: In 1897? In the laboratory? In Germany? Students could legitimately give any one of these three answers. And what if a cagey student wrote the answer "in modern times"—is that right or wrong?

The item could be saved by rephrasing it. Students could be asked specifically in what year or in what country, for example, the cathode ray tube was introduced by Braun, as in the following examples:

In what year was the cathode ray tube with fluorescent screen first introduced? _____

In what country was the cathode ray tube with fluorescent screen first introduced? _____

Don't go overboard with your blanks. This guideline is closely related to the previous one. If

you put too many blanks in one item, there won't be enough significant information left in the item for students to know how to answer it. The following is an example of too many blanks spoiling the broth:

In the _____, _____enters the _____

There really is no way for students to know how to complete this item. However, you could delete the first blank, replacing it with the word it stands for, and improve the item as follows:

In the lungs, _____ enters the _____.

The item is now clear enough that students should be able to supply the correct answers (i.e., *oxygen* enters the *bloodstream*). If you have any question concerning how many blanks you can put in a given item, remember that fewer is usually better. If necessary, you should write several items, each with one blank, to replace a single item with several blanks that is impossible to understand.

Do not turn completion items into essay items. As stated previously, another name for the completion item is the short-answer item. Completion items differ from essay items mainly in the length of the answer required. You should ensure that your completion items are not really essay items (even mini-essay items) in disguise. This point applies particularly to completion items in the form of questions.

Keep blanks in all items the same length. In practice, this means that all blanks should be long enough for the longest answer to be used in the test. You should never leave a short blank for a short answer, medium blank for a medium answer, and so on. Students will quickly understand that there is a clue to the answer in the length of the blank if you do.

Don't give grammatical clues to the answer. Consider the following example:

The _____ remove waste substances from the blood.

Many students will be able to use their knowledge of the English language to help answer this item. The verb *remove* in the item is plural (he, she, or it removes; they remove). Any student who realizes this will also realize that a plural answer is required to be the subject of the verb. Otherwise, the sentence would be grammatically incorrect. This example could be rewritten as follows:

Waste substances are removed from the blood by the _____.

There are no grammatical clues in this version, since we use *the* with both singular and plural nouns in English. You should also avoid having either *a* or *an* immediately before the blank. The indefinite article *a* tells students that the next word starts with a consonant, and *an* indicates that the next word starts with a vowel.

Place the blanks for easy scoring. It is helpful to put all the blanks in one column, in either the left or right margin. Students actually write their answers in this column of blanks. A short blank in the middle of items can indicate which word or phrase students are to supply. The following is an example:

Explosive gas is used as an energy source in the _(3)_ engine. 3. _____

Prepare a scoring key. This can make scoring a lot easier, especially if you put all the blanks in a column. List the correct answers to the items as you write them. Then, as you score the test, you may find other answers that are acceptable in students' papers. You should add these to your key as well.

Assign the same number of points to each blank. Each blank should test student recall of a significant bit of knowledge. Each answer should thus have the same weight, since each item should be of comparable importance and difficulty. This also makes computation of students' scores on the test easier.

Keep all completion items together. This guideline always applies, no matter what type of item is involved. Students should not have to switch from one type of item to another—from one mode of reasoning and answering to another—any more than necessary.

Select one or more cognitive student performance objectives in your occupational specialty that require students to recall information.

Construct five completion items to measure achievement of the student performance objectives you selected. Number each item for easy reference during feedback. Include directions and a scoring key.

You may wish to ask one or more of your peers who have taken or are taking this module to critique the completion items you have constructed. Discuss any suggested changes and then make any necessary revisions.

After you have constructed your completion items, use the Checklist for Completion Items, p. 43, to evaluate your work.

CHECKLIST FOR COMPLETION ITEMS

Directions: Place an X in the NO, PARTIAL, or FULL box to indicate that each of the following performance components was not accomplished, partially accomplished, or fully accomplished. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

Name _____

Date _____

Resource Person _____

	LEVEL OF PERFORMANCE			
The completion items meet the following criteria:	N/A	No	Partial	Full
1. Only significant words are omitted in incomplete sentence items	☐	☐	☐	☐
2. Enough clues are left in the incomplete sentence items so that the required responses make sense to the students	☐	☐	☐	☐
3. There are no grammatical clues to reveal the correct responses	☐	☐	☐	☐
4. All the blanks are the same length	☐	☐	☐	☐
5. All omitted phrases or words are of comparable importance	☐	☐	☐	☐
6. Any question asked is specific enough to evoke the correct type of response	☐	☐	☐	☐
7. Direct quotes from a textbook are avoided	☐	☐	☐	☐
8. All answers are a word, numeral, symbol, or short phrase	☐	☐	☐	☐
9. All answers have equal point value	☐	☐	☐	☐
10. An appropriate scoring key is provided	☐	☐	☐	☐

Level of Performance: All items must receive FULL or N/A responses. If any item received a NO or PARTIAL response, review the material in the information sheet, Completion Items, pp. 40–41, revise your items accordingly, or check with your resource person if necessary.

NOTES

Learning Experience V

OVERVIEW

After completing the required reading, construct ten true-false items to measure the achievement of selected cognitive student performance objectives.

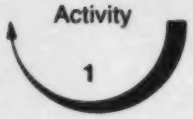
You will be reading the information sheet, True-False Items, pp. 46–48.

You will be selecting one or more student performance objectives that require student knowledge or comprehension of information.

You will be constructing ten true-false items to measure achievement of the student performance objectives you select.

You may wish to ask peers to give suggestions for the improvement of the true-false items you develop.

You will be evaluating your competency in constructing true-false items, using the Checklist for True-False Items, p. 51.



True-false items have perhaps the most definite advantages and disadvantages of the objective forms of testing. For information on how to construct true-false items, read the following information sheet.

TRUE-FALSE ITEMS

True-false items are also called **alternate-response items.** They are selection items in which students must choose one of two given answers. Most commonly, students are asked to judge whether statements are true or false. Students can also be given a list of questions to which they must answer yes or no.

Uses of True-False Items

True-false items can measure student achievement at the knowledge and comprehension levels of the cognitive domain. Furthermore, students can answer true-false items rather quickly, so they can be used to sample a wide range and large amount of content. Finally, true-false items can be scored completely objectively.

Against these advantages, however, you must weigh one large disadvantage. Since students have only two answers to choose from, their chances of guessing correctly are 50–50. Because of this, true-false items are less reliable. They do not discriminate as well between students who know the material and those who don't.

Constructing True-False Items

You can minimize the effects of guessing by following some basic guidelines in writing true-false items. You can avoid giving unnecessary clues that students can use in guessing.

Keep items simple. The longer and more complicated the item is, the more you are testing students' reading ability instead of their technical knowledge. Longer items also give students more opportunities to "read into" the statement an interpretation you didn't intend. Simple items can be more clearly true or false.

Furthermore, many teachers tend to write true items that are consistently longer than false items. This is because, often, a statement must be qualified or modified to be sure it really is true. Experienced test-takers soon realize that a long item is likely to be true, thereby improving their chances of guessing correctly.

Make each item entirely true or entirely false. Sometimes teachers write items that are partly true, partly false. Consider the following example:

> *T F The rip saw is best suited to cutting along the grain because of the angle of the grip to the blade.*

The student who knows his or her rip saws will realize that part of this statement is true—rip saws are used to cut along the grain. However, the student will also know that this has nothing to do with grip. Is this statement true for the cut or false for the grip?

Don't make items false on a technicality. Closely related to the last guideline, this rule means that you should test for **significant** bits of knowledge. Consider the following example:

> *T F Pasteur first inoculated a human, a young child, with his treatment of rabies on July 7, 1885.*

The answer is false—the inoculation occurred on July 6. The item, then, becomes trivial. It is difficult to imagine why students would need to know the exact day of the inoculation. It would be equally trivial to change the year to 1886. You will need to use your own knowledge of technical content to determine when details are significant and when they are not.

Avoid qualitative terms in your items. When you are writing items, don't use words that express indefinite quantities (e.g., *several*, *some*, *many*). These make the statement open to different interpretations, as in the following example:

> *T F Several different thickeners can be used in making gravy.*

If *several* means two or three, this item would be true. But if *several* means five or six, the item would be false. You could improve the item by using a specific number instead of the word *several.*

Quote authorities for opinions. You might use true-false items to test students' knowledge of opinion in your area. It may be significant that established authorities in the field believe that a technological development will become widespread. But you

must tell students whose opinion you are quoting, so they can know how to answer. The following example would tell students what they need to know:

> *T F The editors of* Keyboard and Console *predict that word processing equipment will replace current technology in office operations within five years.*

Avoid negative items. Negative statements are more difficult for students to read and consequently favor students with higher reading skills. You should avoid items with negative words—*not*, *never*, *nothing*, *no*, and so on. You should especially avoid double negatives, as in the following example:

> *T F Broken ears are a not uncommon defining characteristic of breeds of swine.*

If students are unable to decode the double negative *not uncommon*, it won't matter whether they know the information you are testing for. The item could be improved by stating it in positive terms, as follows:

> *T F Broken ears are a common defining characteristic of breeds of swine.*

Avoid specific determiners. Certain words are usually associated with either true or false statements. Statements are usually false if they contain the words *always*, *never*, *all*, or *none*. On the other hand, the words *sometimes*, *some*, *may*, and *should* generally appear in true statements. If you use these words in your items, you are giving students an unnecessary clue to the answer.

Use equal numbers of true and false statements. While you don't need to take this point literally, you should have **about** the same number of true and false statements. If you consistently have more of one kind than the other, students can use this as a clue in guessing. When you have finished writing all your items, you can count your true and false items and add items as necessary to make them roughly equal.

Prepare a scoring key. Scoring is very easy if you prepare a key. The best would be a stencil—a copy of the test with the correct answers punched out of the *T* and *F* columns. You can place this over students' papers and see their answers at a glance. If you use a stencil, you should have students mark their answers by placing an *X* over the *T* or *F*. If they circle their answers, they could make their circles bigger than the holes in your stencil and you wouldn't see them.

Give simple and clear directions. You should always tell students clearly and simply what they are to do. Your directions should include how students are to mark their answers (e.g., place an *X* over the *T* or *F*). The following is an example of directions for true-false items:

> **Directions:** Each of the following statements is either true or false. Read each statement and decide whether it is true or false. Indicate your answer by placing an *X* over the *T* if the statement is true and over the *F* if the statement is false.

Watch out for patterns of answers. If you always use two true items followed by two false items, students will soon recognize this pattern. You should review the answers when you have finished writing your items. If you notice any pattern in the answers, you can use various procedures to distribute the answers randomly (e.g., rolling dice or pulling numbered slips out of a hat).

Don't quote from the textbook. This practice only places a premium on rote learning and does not encourage higher-level cognitive activities. In addition, many textbook statements are ambiguous when removed from their original context.

Consider using modified true-false items. Because responding to regular true-false items often becomes a kind of guessing game for students, modified true-false items are often used when recognition **and** recall of information are being measured. There are many methods for modifying true-false items. The main disadvantage is in the grading of the items, in that more time is required.

The modification consists of asking students to correct or give the right response to any false item. The value in the modification is that the students are required not only to recognize a false statement, but to recall the **correct** response as well.

Two forms of this modification are illustrated in samples 5 and 6. Notice that in both of these modifications, the word or words that make the statement true or false are underlined. In sample 5, students must supply their own correct response. In sample 6, each true-false item provides multiple choices for correcting the item if it is false. These are much more difficult to construct, because you have to list the plausible choices for each true-false item.

In the form illustrated in sample 6, grading or giving of points for the responses could be a problem. One suggestion is to give each item, whether true or false, a value of 2. Recognition of a true statement is rewarded with 2 points; a value of 1 is given for recognizing a false statement; and a value of 1 is given for recalling the correct response.

SAMPLE 5

MODIFIED TRUE-FALSE ITEM

Directions: Circle the T or F in the column to the left of each true (T) or false (F) statement. **Then,** if the statement is false, correct the statement in the space provided under each statement.

T F 1. The horizontal threads in fabric are called the warp threads.

Correct statement: _____

_____

SAMPLE 6

MODIFIED TRUE-FALSE ITEM

Directions: Circle the T or F in the column to the left of each true (T) or false (F) statement. **Then,** if the statement is false, select the correct alternative and place the letter in the space provided in the left-hand column.

T F 1. The primary colors are blue, green, and yellow.

_____

a. green, orange, and red
b. green, red, and yellow
c. violet, yellow, and red
d. orange, blue, and violet
e. yellow, green, and violet

Select one or more cognitive student performance objectives in your occupational specialty that require student knowledge or comprehension of information.

Construct ten true-false items to measure achievement of the student performance objectives you selected. Number each item for easy reference during feedback. Include directions and a scoring key.

You may wish to ask one or more of your peers who have taken or are taking this module to critique the true-false items you have constructed. Discuss any suggested changes and then make any necessary revisions.

After you have constructed your true-false items, use the Checklist for True-False Items, p. 51, to evaluate your work.

NOTES

CHECKLIST FOR TRUE-FALSE ITEMS

Directions: Place an X in the NO, PARTIAL, or FULL box to indicate that each of the following performance components was not accomplished, partially accomplished, or fully accomplished. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

Name _____

Date _____

Resource Person _____

	LEVEL OF PERFORMANCE			
	N/A	No	Partial	Full
The true-false items meet the following criteria:				
1. All statements are entirely true or entirely false	☐	☐	☐	☐
2. No trivial details are used to make any of the statements false	☐	☐	☐	☐
3. The statements are concise without more elaboration than is necessary to give clear meaning	☐	☐	☐	☐
4. The statements have not been quoted from a textbook	☐	☐	☐	☐
5. The statements avoid using specific determiners that signal answers	☐	☐	☐	☐
6. The items are stated in positive rather than negative terms	☐	☐	☐	☐
7. An authority is quoted when the item contains opinion	☐	☐	☐	☐
8. A pattern of responses is avoided in the items	☐	☐	☐	☐
9. The responses for the group of items are written in a manner that makes scoring easy (e.g., in a column to the left)	☐	☐	☐	☐
10. Statements avoid qualitative terms (e.g., *several*, *many*)	☐	☐	☐	☐
11. Numbers of true and false items are roughly equal	☐	☐	☐	☐
12. Clear and simple directions are provided	☐	☐	☐	☐
13. An appropriate scoring key is provided	☐	☐	☐	☐

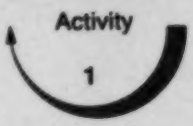
Level of Performance: All items must receive FULL or N/A responses. If any item receives a NO or PARTIAL response, review the material in the information sheet, True-False Items, pp. 46–48, revise your items accordingly, or check with your resource person if necessary.

NOTES

Learning Experience VI

OVERVIEW

After completing the required reading, construct three essay items and two oral items to measure the achievement of selected cognitive student performance objectives.

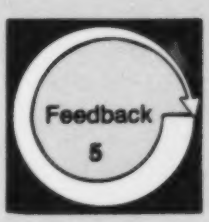


You will be reading the information sheet, Essay and Oral Items, pp. 54-55.

You will be selecting one or more student performance objectives that require students to apply, analyze, synthesize, or evaluate information.

You will be constructing three essay items and two oral items to measure achievement of the student performance objectives you select.

You may wish to ask peers to give suggestions for the improvement of the essay and oral items you develop.



You will be evaluating your competency in constructing essay and oral items, using the Checklist for Essay and Oral Items, p. 57.

Essay and oral items have distinct advantages and disadvantages. Properly constructed, they can be excellent evaluation devices. For information on writing high-quality essay and oral items, read the following information sheet.

ESSAY AND ORAL ITEMS

Essay and oral items, like completion items, are **supply** items. Completion items, however, require students to supply their own knowledge one piece at a time. Essay or oral items require students to recall and supply **numerous** pieces of knowledge. Furthermore, they require students to organize and present this knowledge in a **logical** and **understandable** manner.

Both essay and oral items can be divided into two kinds, depending on the kind of response they call for. The **extended-response item** allows students considerable freedom in structuring and presenting their own answers. The following item, for example, could be answered in a number of different ways:

> ***Discuss the stand of the American Medical Association on socialized medicine.***

A **restricted-response item,** on the other hand, limits the content and form of the response in some way. Both the item and the response it calls for are much more specific, as in the following example:

> ***Describe the criticisms made by the American Medical Association against socialized medicine. Present one example to support each criticism and one example to counter each criticism.***

Uses of Essay and Oral Items

Essay and oral items can be used to test student performance at any level in the cognitive domain. They are most useful in evaluating performance on high-level instructional objectives requiring the application, analysis, synthesis, or evaluation of information.

Compared to objective items, there are some "disadvantages" in using essay and oral items. Scoring of students' answers to essay and oral items is more **subjective** and, thus, can be more unreliable. Even a restricted-response item allows students some latitude in responding, and scoring must allow equal credit for different responses of equal merit. Thus, it is critical that you prepare and utilize a structured, yet flexible, answer key in scoring responses, so that your ratings are as objective and fair as possible.

Answering essay and oral items and scoring the answers also demand a good deal of **time.** Oral tests must be conducted and scored individually, which may require large amounts of teacher and student time. Teachers can often expect to spend several hours scoring essay items for a whole class—unlike objective items, which can usually be answered and scored in a matter of minutes.

Finally, these items test more than just students' knowledge of technical content. They also test students' **ability to communicate,** orally or in writing. Students must be able not only to recall and apply, analyze, synthesize, or evaluate technical knowledge. They must also be able to communicate their knowledge and their treatment of it to you.

However, when the student performance objectives call for higher levels of cognition and when ability to organize and communicate information is important for occupational success, then essay and oral items should be used. The time involved in scoring will be well worth it. Short-answer tests are useful and have their place in evaluation, but only through essay and oral items can you test, and challenge, students' ability to think—to take what they know, analyze it, synthesize it, organize it, and communicate their ideas to others.

Constructing Essay and Oral Items

There are several guidelines that you can follow to minimize any limitations of essay and oral items.

Develop clear, precise items. Your essay and oral items should always communicate your intent to students with no room for misinterpretation. In practice, restricted-response items are usually preferable. Each item should tell students, as clearly as possible, what information is required in the answer and how that information should be presented.[4]

Clear, precise items allow less opportunity for students to misunderstand an item. They also reduce the opportunities for students to evade responding to an item by speaking or writing at length in generalities. Scoring also becomes easier when the required content and form of the response are clearly indicated in the item.

Require students to answer all items. Teachers sometimes make some essay or oral items optional.

4. To gain skill in presenting questions orally to students, you may wish to refer to Module C-12, *Employ Oral Questioning Techniques.*

Students might be directed, for example, to respond to any one of three items. However, when different students answer different items, it is almost impossible to get comparable scores because students have not been performing comparable tasks. With optional items, you end up testing different groups of students on different bodies of information and different ways of treating those different bodies of information.

Use a larger number of shorter items. Since students take more time in responding to essay or oral items, you cannot sample as wide a range of knowledge as with objective items. The best way to make your sample as wide as possible is to use a larger number of items that require shorter answers. Your essay or oral items will then evaluate a wider variety of instructional objectives.

Develop items for higher-level objectives and content. Given the nature of essay and oral items, you should use them to measure only appropriate, high-level instructional achievement. Objective items should be used to measure the lower cognitive levels (i.e., knowledge and comprehension)—a task they can accomplish accurately and efficiently. The use of each type of item—subjective and objective—should be restricted to the levels of accomplishment for which each is most suited.

Provide full and clear directions. In your directions, you should tell students exactly what their task is—for example, to answer each of the three items provided on a written essay test. You should tell students what the point value of each item is and how much time is available for each item or for the whole test. Students will then know approximately how much time to spend on each item. Students must also be told if their scores will be affected by mechanical factors (e.g., spelling and handwriting on essay items, grammar on essay or oral items).

Give students advance notice and practice tests. You should give your students warning of upcoming essay or oral tests. This provides them with an opportunity to review the material to be covered from the appropriate perspective. You probably know from your own experiences that preparing for an essay test requires the use of different study skills than preparing for objective tests.

Furthermore, students should be given practice tests. Some students may never have been required to respond to essay or oral items before. They may have had no experience in organizing and presenting a body of information, as these items require. They may not be aware that they will need to spend some of the time allotted for the test in planning their answers. Students should have the chance to practice answering oral or essay items before being evaluated on their performance of these skills.

Don't use open-book tests. Allowing students to use their books doesn't really help them to answer items—unless you are also testing their ability to locate appropriate information in reference books. Remember that an essay or oral item should be used to evaluate student ability at the higher cognitive levels. At these levels, it matters somewhat less what knowledge students supply. More important is how they can apply that knowledge in answering the essay or oral item. Studies have shown that students don't do any better or worse than expected when they are allowed to use their books.

Prepare a structured key for scoring. The subjectivity of scoring can be minimized if you write a structured scoring key—a model answer to the essay or oral item, usually in outline form. As you write your own answer to the item, you can assign specific point values to specific aspects of the answer. Students' answers can then be compared to the key to assign point values to their answers.

It is a good idea to have your scoring key reviewed by a colleague who is also knowledgeable in your occupational area. He or she may identify other information to include in your model answer. Likewise, you may find information in students' answers that is correct but that you did not anticipate. Students might include facts not mentioned in your key, or they might organize and present their facts in a way you hadn't thought of. Any such information should be added to your key as well.

Read all students' answers to one essay item at one sitting. The best way to score answers to essay items consistently is to read all students' answers to one item at one sitting. In other words, don't read one student's entire test before reading another's. You can review your key for the item you are scoring and concentrate on that item. If possible, you should not even know whose answer you are reading. Studies have shown that the "halo effect" is real—teachers tend to score answers higher for students who are known to be more capable.

Score mechanical factors separately. Experts disagree on whether such factors as handwriting, spelling, or grammar should be considered at all in scoring students' answers. Some are all for it; some are equally against it. You will need to use your own judgment to determine whether to score students' answers for these factors. If these skills are important for success in the occupation, then they should, at some point and in some way, be evaluated.

If you do score for these factors, however, you should assign a score separate from that for technical content. Regardless of how important skill in communicating may be in your occupational area, the technical content of students' answers should be scored on its own merit.

Select one or more student performance objectives in your occupational specialty that require students to use the higher cognitive levels (application, analysis, synthesis, evaluation).

Construct three essay items and two oral items to measure achievement of the student performance objectives you selected. Number each item for easy reference during feedback. Include directions and a scoring key.

You may wish to ask one or more of your peers who have taken or are taking this module to critique the essay and oral items you have constructed. Discuss any suggested changes and then make any necessary revisions.

After you have constructed your essay and oral items, use the Checklist for Essay and Oral Items, p. 57, to evaluate your work.

CHECKLIST FOR ESSAY AND ORAL ITEMS

Directions: Place an X in the NO, PARTIAL, or FULL box to indicate that each of the following performance components was not accomplished, partially accomplished, or fully accomplished. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

Name _____

Date _____

Resource Person _____

	LEVEL OF PERFORMANCE			
The essay and oral items meet the following criteria:	N/A	No	Partial	Full
1. The language in each item is precise and unambiguous	☐	☐	☐	☐
2. The problems or tasks require students to use the higher levels of cognition	☐	☐	☐	☐
3. Open books are not allowed	☐	☐	☐	☐
4. Directions clearly state the task, point value of items, and time allowed for answering	☐	☐	☐	☐
5. The items clarify any additional directions needed beyond the general set of directions	☐	☐	☐	☐
6. Students are required to answer all items	☐	☐	☐	☐
7. A structured scoring key is provided	☐	☐	☐	☐

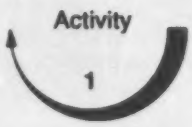
Level of Performance: All items must receive FULL or N/A responses. If any item receives a NO or PARTIAL response, review the material in the information sheet, Essay and Oral Items, pp. 54–55, revise your items accordingly, or check with your resource person if necessary.

NOTES

Learning Experience VII

OVERVIEW

After completing the required reading, construct one case study or problem-solving item to measure the achievement of a selected cognitive student performance objective.



You will be reading the information sheet, Case Studies and Problem-Solving Items, pp. 60–61.

You will be selecting one student performance objective that requires students to apply, analyze, synthesize, or evaluate information.

You will be constructing one case study or problem-solving item to measure achievement of the student performance objective you select.

You may wish to ask peers to give suggestions to improve the clarity of the case study or problem-solving item you develop.

You will be evaluating your competency in constructing a case study or problem-solving item, using the Checklist for Case Studies and Problem-Solving Items, p. 63.

The case study or problem-solving technique is often used as a method of instruction. For information on how to use these techniques in evaluating student knowledge, read the following information sheet.

CASE STUDIES AND PROBLEM-SOLVING ITEMS

Case studies and problem-solving items present students with a situation to which they must respond. These items differ from other types of test items in the amount of detail they contain. Objective items are usually short—one sentence. Essay or oral items are usually no longer than three or four sentences. Some case studies or problem-solving items can be much longer; each item might be a page or more in length.

Case studies and problem-solving items are usually presented to students in written form. You might, for example, write a case study describing how a worker observed safety regulations on the job. Students could read the case study and critique the worker's safety practices. Students could also be given a table, graph, or diagram and asked to answer items on the basis of it.

But such items do not necessarily have to be in written form. Students might be asked to react to case studies or problem-solving items presented on videotape or audiotape.

Students are asked to use the information presented in the case study or problem-solving item to analyze a situation or solve a problem. Often, other knowledge is used as well. Students in health care occupations, for instance, could be given a series of problem-solving items, each of which describes a patient's symptoms in a particular emergency condition (e.g., choking on a piece of food). Students could be asked to identify the condition and describe the correct emergency procedures to follow.

Students can also be asked to respond to the situation or solve the problem in a number of different ways. They might answer essay or oral items in which they evaluate methods, materials or solutions. True-false items could be used to evaluate student comprehension of the data presented in a case study. Students' ability to troubleshoot a malfunctioning machine could be evaluated using multiple-choice items.

Uses of Case Studies and Problem-Solving Items

The case study or problem-solving statement could ask students to **predict trends** or to **make inferences of consequences** (high-level types of knowing). For example, assume you are teaching students the factors to consider in selecting an occupation. For evaluation purposes you might want to design a case study in which students would have to predict from the data presented which jobs will be available in the next two years.

You can also use case studies and problem-solving items when the objective is to teach the **application** of principles, generalizations, and concepts. For example, an instructor could ask the students what automotive repair principles they would use to solve a list of problem situations faced in an auto body shop. This type of item requires students to **apply** what they have learned.

Some objectives stress the ability of students to **analyze**—to break down material into its parts, detect relationships of the parts, and recognize the organizational principles of the structure as a whole. These require a higher level of knowing, and case studies and problem-solving items are frequently used for evaluating student performance at this level.

For example, if a teacher wanted students in office machine repair to be able to troubleshoot (analyze) a defective calculating machine, he or she could develop a problem-solving item listing all the symptoms of a malfunctioning calculator and then ask students to describe the probable cause of the trouble.

The case study or problem-solving item is also an excellent means of evaluating students' ability to **judge** (**evaluate** or **appraise**) ideas, methods, materials, solutions, and so on. For example, achievement of objectives involving the purchasing of goods

and services can be evaluated through case studies, in that such objectives require students to make value judgments.

Constructing Case Studies and Problem-Solving Items

Several guidelines can help you to write clear case studies and problem-solving items. In general, your items should be practical and realistic and contain enough detail for students to perform the task required in relation to the case study or problem.

Describe a realistic situation. A marketing and distributive education instructor, for example, could write a case study describing how a retail clerk waited on customers and ask students to critique the clerk's customer service. The clerk's performance should be realistic—neither all good nor all bad, since people are neither one nor the other. Also, mistakes in performance should not be obvious. Don't, for example, have the retail clerk chase a customer out of the store.

Give students all the detail they will need for the task. For example, a case study might list the symptoms of a malfunctioning carburetor. Students could be asked to identify probable causes of the trouble. They could even be asked to describe how they would verify their diagnoses. To complete this task successfully, they would need to know all the relevant detail from reading the case study. If you expect students to identify a particular cause for a problem, be sure they have enough information to come to that conclusion.

You can still hold your students responsible for other knowledge in the task, of course. If a case study involves the application of Ohm's Law to an electronic circuit design, the instructor could legitimately expect students to know what Ohm's Law is and how to apply it (assuming they have studied those things already). That, in fact, would be one of the things the instructor would be testing in this item. He or she would have to be sure, however, to supply in the case study all the information students would need on the circuit design.

Clearly state the task to be performed. Students should be told clearly what they are to do with the case study or problem once they have read it. If there are objective items for students to answer, you need to tell them to answer the items using the information in the case study or problem statement. If students are to write longer, essay-type responses (e.g., analyzing, critiquing, evaluating), you need to tell them how detailed their answers should be.

Follow good item-writing guidelines. As stated previously, you might ask students to answer true-false items about a case study, multiple-choice items about a circuit diagram, or essay items evaluating a worker's safety practices. In each case, you should follow all the guidelines for writing that type of item. Specifically, you should provide directions for answering the items. The type of item you use should be chosen to fit the learning activity being evaluated.

Write items in terms students can understand. Your case studies or problem-solving items should not penalize students with lower reading or oral communication skills. You should use technical terms that students know or define those terms in your items.

Use appropriate scoring keys and procedures. You should prepare scoring keys (e.g., outlines of model answers) for your case studies or problem-solving items. Likewise, you should use appropriate procedures for scoring student responses to the items you have used.

Select a student performance objective in your occupational specialty that requires students to apply, analyze, synthesize, or evaluate information.

Construct one case study or problem-solving item to measure achievement of the student performance objective you selected. Include directions and a scoring key.

You may wish to ask one or more of your peers who have taken or are taking this module to critique the case study or problem-solving item you have constructed. Discuss any suggested changes and then make any necessary revisions.

After you have constructed your case study or problem-solving item, use the Checklist for Case Studies and Problem-Solving Items, p. 63, to evaluate your work.

CHECKLIST FOR CASE STUDIES AND PROBLEM-SOLVING ITEMS

Directions: Place an X in the NO, PARTIAL, or FULL box to indicate that each of the following performance components was not accomplished, partially accomplished, or fully accomplished. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

Name _____

Date _____

Resource Person _____

The case study or problem-solving item meets the following criteria:	LEVEL OF PERFORMANCE			
	N/A	No	Partial	Full
1. The situation developed is based on the student performance objective	☐	☐	☐	☐
2. The situation is practical and realistic	☐	☐	☐	☐
3. The situation is described in sufficient detail for the task	☐	☐	☐	☐
4. The problem or case study is worded at the students' comprehension level	☐	☐	☐	☐
5. The task requires students to apply, analyze, synthesize, or evaluate learning	☐	☐	☐	☐
6. The item contains a clear statement of what students are to do	☐	☐	☐	☐
7. The item indicates the level of detail required in the solution	☐	☐	☐	☐
8. Test items concerning the case study or problem follow good item-writing guidelines	☐	☐	☐	☐
9. Item directions state point values and time limits	☐	☐	☐	☐
10. Appropriate scoring keys are provided	☐	☐	☐	☐

Level of Performance: All items must receive FULL or N/A responses. If any item receives a NO or PARTIAL response, review the material in the information sheet, Case Studies and Problem-Solving Items, pp. 60–61, revise your items accordingly, or check with your resource person if necessary.

NOTES

Learning Experience VIII

FINAL EXPERIENCE

In an **actual teaching situation,*** assess student cognitive (knowledge) performance.

As you fulfill your teaching duties, determine when the nature of the student performance objectives requires the assessment of student knowledge. Based on that decision, assess student cognitive performance. This will include—

- determining the appropriate types of test items to use to measure students' achievement of unit and lesson objectives
- constructing at least one unit test designed to measure student progress toward the objectives
- administering the unit test to students

NOTE: As you complete each of the above activities, document your actions (in writing, on tape, through a log) for assessment purposes.

Arrange to have your resource person review your test and other documentation (e.g., audiotape of your directions to students prior to administering the test).

Your total competency will be assessed by your resource person, using the Teacher Performance Assessment Form, pp. 67–68.

Based upon the criteria specified in this assessment instrument, your resource person will determine whether you are competent in assessing student cognitive (knowledge) performance.

* For a definition of "actual teaching situation," see the inside back cover.

NOTES

TEACHER PERFORMANCE ASSESSMENT FORM

Assess Student Performance: Knowledge (D–2)

Name _____

Date _____

Resource Person _____

Directions: Indicate the level of the teacher's accomplishment by placing an X in the appropriate box under the LEVEL OF PERFORMANCE heading. If, because of special circumstances, a performance component was not applicable, or impossible to execute, place an X in the N/A box.

	LEVEL OF PERFORMANCE					
	N/A	None	Poor	Fair	Good	Excellent
In developing a unit test to assess student cognitive performance, the teacher:						
1. identified objectives in the cognitive (knowledge) area	☐	☐	☐	☐	☐	☐
2. identified types of test items appropriate for assessing the level of knowledge implied in the performance objectives	☐	☐	☐	☐	☐	☐
3. developed each test item according to the guidelines for constructing that type of item	☐	☐	☐	☐	☐	☐
4. constructed a test that was:						
a. valid	☐	☐	☐	☐	☐	☐
b. reliable	☐	☐	☐	☐	☐	☐
c. objective	☐	☐	☐	☐	☐	☐
d. discriminating	☐	☐	☐	☐	☐	☐
e. comprehensive	☐	☐	☐	☐	☐	☐
f. easy to use	☐	☐	☐	☐	☐	☐
5. grouped together all test items of like type	☐	☐	☐	☐	☐	☐
6. grouped together within each type the items relating to the same objective	☐	☐	☐	☐	☐	☐
7. included no more than three or four different types of test items	☐	☐	☐	☐	☐	☐
8. sequenced like items logically	☐	☐	☐	☐	☐	☐
9. included clear and concise directions for each different type of test item	☐	☐	☐	☐	☐	☐
10. constructed the test so that recording responses was simple and consistent	☐	☐	☐	☐	☐	☐
11. numbered items consecutively from the beginning to the end of the test	☐	☐	☐	☐	☐	☐
12. developed an appropriate key for scoring	☐	☐	☐	☐	☐	☐

In administering the test, the teacher:	N/A	None	Poor	Fair	Good	Excellent
13. arranged the physical environment so it was suitable for test-taking	☐	☐	☐	☐	☐	☐
14. explained the use of the test to students in a nonthreatening way	☐	☐	☐	☐	☐	☐
15. informed students concerning how the test would be graded prior to their taking the test	☐	☐	☐	☐	☐	☐
16. gave specific instructions for taking the test	☐	☐	☐	☐	☐	☐
17. had all needed testing materials and supplies ready for distribution	☐	☐	☐	☐	☐	☐
18. indicated approximate time needed to complete the test by sections and as a whole	☐	☐	☐	☐	☐	☐
19. allowed sufficient time for the test to be completed by all or most students	☐	☐	☐	☐	☐	☐
20. answered students' questions during the test in a manner that did not disturb the total group	☐	☐	☐	☐	☐	☐

Level of Performance: All items must receive N/A, GOOD, or EXCELLENT responses. If any item receives a NONE, POOR, or FAIR response, the teacher and resource person should meet to determine what additional activities the teacher needs to complete in order to reach competency in the weak area(s).

ABOUT USING THE NATIONAL CENTER'S PBTE MODULES

Organization

Each module is designed to help you gain competency in a particular skill area considered important to teaching success. A module is made up of a series of learning experiences, some providing background information, some providing practice experiences, and others combining these two functions. Completing these experiences should **enable** you to achieve the **terminal** objective in the final learning experience. The final experience in each module always requires you to demonstrate the skill in an actual teaching situation when you are an intern, a student teacher, an inservice teacher, or occupational trainer.

Procedures

Modules are designed to allow you to individualize your teacher education program. You need to take only those modules covering skills that you do not already possess. Similarly, you need not complete any learning experience within a module if you already have the skill needed to complete it. Therefore, before taking any module, you should carefully review (1) the introduction, (2) the objectives listed on p. 4, (3) the overviews preceding each learning experience, and (4) the final experience. After comparing your present needs and competencies with the information you have read in these sections, you should be ready to make one of the following decisions:

- That you do not have the competencies indicated and should complete the entire module
- That you are competent in one or more of the enabling objectives leading to the final learning experience and, thus, can omit those learning experiences
- That you are already competent in this area and are ready to complete the final learning experience in order to "test out"
- That the module is inappropriate to your needs at this time

When you are ready to complete the final learning experience **and** have access to an actual teaching situation, make the necessary arrangements with your resource person. If you do not complete the final experience successfully, meet with your resource person and arrange to (1) repeat the experience or (2) complete (or review) previous sections of the module or other related activities suggested by your resource person before attempting to repeat the final experience.

Options for recycling are also available in each of the learning experiences preceding the final experience. Any time you do not meet the minimum level of performance required to meet an objective, you and your resource person may meet to select activities to help you reach competency. This could involve (1) completing parts of the module previously skipped, (2) repeating activities, (3) reading supplementary resources or completing additional activities suggested by the resource person, (4) designing your own learning experience, or (5) completing some other activity suggested by you or your resource person.

Terminology

Actual Teaching Situation: A situation in which you are actually working with and responsible for teaching secondary or postsecondary vocational students or other occupational trainees. An intern, a student teacher, an inservice teacher, or other occupational trainer would be functioning in an actual teaching situation. If you do **not** have access to an actual teaching situation when you are taking the module, you can complete the module **up to** the final learning experience. You would then complete the final learning experience later (i.e., when you have access to an actual teaching situation).

Alternate Activity or Feedback: An item that may **substitute** for required items that, due to special circumstances, you are unable to complete.

Occupational Specialty: A specific area of preparation within a vocational service area (e.g., the service area Trade and Industrial Education includes occupational specialties such as automobile mechanics, welding, and electricity.

Optional Activity or Feedback: An item that is not required but that is designed to **supplement** and enrich the required items in a learning experience.

Resource Person: The person in charge of your educational program (e.g., the professor, instructor, administrator, instructional supervisor, cooperating/supervising/classroom teacher, or training supervisor who is guiding you in completing this module).

Student: The person who is receiving occupational instruction in a secondary, postsecondary, or other training program.

Vocational Service Area: A major vocational field: agricultural education, business and office education, marketing and distributive education, health occupations education, home economics education, industrial arts education, technical education, or trade and industrial education.

You or the Teacher/Instructor: The person who is completing the module.

Levels of Performance for Final Assessment

N/A: The criterion was not met because it was **not applicable** to the situation.

None: No attempt was made to meet the criterion, although it was relevant.

Poor: The teacher is unable to perform this skill or has only **very limited ability** to perform it.

Fair: The teacher is unable to perform this skill in an acceptable manner but has **some ability** to perform it.

Good: The teacher is able to perform this skill in an **effective** manner.

Excellent: The teacher is able to perform this skill in a **very effective** manner.

Titles of the National Center's Performance-Based Teacher Education Modules

Category A: Program Planning, Development, and Evaluation

A–1 Prepare for a Community Survey
A–2 Conduct a Community Survey
A–3 Report the Findings of a Community Survey
A–4 Organize an Occupational Advisory Committee
A–5 Maintain an Occupational Advisory Committee
A–6 Develop Program Goals and Objectives
A–7 Conduct an Occupational Analysis
A–8 Develop a Course of Study
A–9 Develop Long-Range Program Plans
A–10 Conduct a Student Follow-Up Study
A–11 Evaluate Your Vocational Program

Category B: Instructional Planning

B–1 Determine Needs and Interests of Students
B–2 Develop Student Performance Objectives
B–3 Develop a Unit of Instruction
B–4 Develop a Lesson Plan
B–5 Select Student Instructional Materials
B–6 Prepare Teacher-Made Instructional Materials

Category C: Instructional Execution

C–1 Direct Field Trips
C–2 Conduct Group Discussions, Panel Discussions, and Symposiums
C–3 Employ Brainstorming, Buzz Group, and Question Box Techniques
C–4 Direct Students in Instructing Other Students
C–5 Employ Simulation Techniques
C–6 Guide Student Study
C–7 Direct Student Laboratory Experience
C–8 Direct Students in Applying Problem-Solving Techniques
C–9 Employ the Project Method
C–10 Introduce a Lesson
C–11 Summarize a Lesson
C–12 Employ Oral Questioning Techniques
C–13 Employ Reinforcement Techniques
C–14 Provide Instruction for Slower and More Capable Learners
C–15 Present an Illustrated Talk
C–16 Demonstrate a Manipulative Skill
C–17 Demonstrate a Concept or Principle
C–18 Individualize Instruction
C–19 Employ the Team Teaching Approach
C–20 Use Subject Matter Experts to Present Information
C–21 Prepare Bulletin Boards and Exhibits
C–22 Present Information with Models, Real Objects, and Flannel Boards
C–23 Present Information with Overhead and Opaque Materials
C–24 Present Information with Filmstrips and Slides
C–25 Present Information with Films
C–26 Present Information with Audio Recordings
C–27 Present Information with Televised and Videotaped Materials
C–28 Employ Programmed Instruction
C–29 Present Information with the Chalkboard and Flip Chart
C–30 Provide for Students' Learning Styles

Category D: Instructional Evaluation

D–1 Establish Student Performance Criteria
D–2 Assess Student Performance: Knowledge
D–3 Assess Student Performance: Attitudes
D–4 Assess Student Performance: Skills
D–5 Determine Student Grades
D–6 Evaluate Your Instructional Effectiveness

Category E: Instructional Management

E–1 Project Instructional Resource Needs
E–2 Manage Your Budgeting and Reporting Responsibilities
E–3 Arrange for Improvement of Your Vocational Facilities
E–4 Maintain a Filing System
E–5 Provide for Student Safety
E–6 Provide for the First Aid Needs of Students
E–7 Assist Students in Developing Self-Discipline
E–8 Organize the Vocational Laboratory
E–9 Manage the Vocational Laboratory
E–10 Combat Problems of Student Chemical Use

Category F: Guidance

F–1 Gather Student Data Using Formal Data-Collection Techniques
F–2 Gather Student Data Through Personal Contacts
F–3 Use Conferences to Help Meet Student Needs
F–4 Provide Information on Educational and Career Opportunities
F–5 Assist Students in Applying for Employment or Further Education

Category G: School–Community Relations

G–1 Develop a School–Community Relations Plan for Your Vocational Program
G–2 Give Presentations to Promote Your Vocational Program
G–3 Develop Brochures to Promote Your Vocational Program
G–4 Prepare Displays to Promote Your Vocational Program
G–5 Prepare News Releases and Articles Concerning Your Vocational Program
G–6 Arrange for Television and Radio Presentations Concerning Your Vocational Program
G–7 Conduct an Open House
G–8 Work with Members of the Community
G–9 Work with State and Local Educators
G–10 Obtain Feedback about Your Vocational Program

Category H: Vocational Student Organization

H–1 Develop a Personal Philosophy Concerning Vocational Student Organizations
H–2 Establish a Vocational Student Organization
H–3 Prepare Vocational Student Organization Members for Leadership Roles
H–4 Assist Vocational Student Organization Members in Developing and Financing a Yearly Program of Activities
H–5 Supervise Activities of the Vocational Student Organization
H–6 Guide Participation in Vocational Student Organization Contests

Category I: Professional Role and Development

I–1 Keep Up to Date Professionally
I–2 Serve Your Teaching Profession
I–3 Develop an Active Personal Philosophy of Education
I–4 Serve the School and Community
I–5 Obtain a Suitable Teaching Position
I–6 Provide Laboratory Experiences for Prospective Teachers
I–7 Plan the Student Teaching Experience
I–8 Supervise Student Teachers

Category J: Coordination of Cooperative Education

J–1 Establish Guidelines for Your Cooperative Vocational Program
J–2 Manage the Attendance, Transfers, and Terminations of Co-Op Students
J–3 Enroll Students in Your Co-Op Program
J–4 Secure Training Stations for Your Co-Op Program
J–5 Place Co-Op Students on the Job
J–6 Develop the Training Ability of On-the-Job Instructors
J–7 Coordinate On-the-Job Instruction
J–8 Evaluate Co-Op Students' On-the-Job Performance
J–9 Prepare for Students' Related Instruction
J–10 Supervise an Employer–Employee Appreciation Event

Category K: Implementing Competency-Based Education (CBE)

K–1 Prepare Yourself for CBE
K–2 Organize the Content for a CBE Program
K–3 Organize Your Class and Lab to Install CBE
K–4 Provide Instructional Materials for CBE
K–5 Manage the Daily Routines of Your CBE Program
K–6 Guide Your Students Through the CBE Program

Category L: Serving Students with Special/Exceptional Needs

L–1 Prepare Yourself to Serve Exceptional Students
L–2 Identify and Diagnose Exceptional Students
L–3 Plan Instruction for Exceptional Students
L–4 Provide Appropriate Instructional Materials for Exceptional Students
L–5 Modify the Learning Environment for Exceptional Students
L–6 Promote Peer Acceptance of Exceptional Students
L–7 Use Instructional Techniques to Meet the Needs of Exceptional Students
L–8 Improve Your Communication Skills
L–9 Assess the Progress of Exceptional Students
L–10 Counsel Exceptional Students with Personal–Social Problems
L–11 Assist Exceptional Students in Developing Career Planning Skills
L–12 Prepare Exceptional Students for Employability
L–13 Promote Your Vocational Program with Exceptional Students

Category M: Assisting Students in Improving Their Basic Skills

M–1 Assist Students in Achieving Basic Reading Skills
M–2 Assist Students in Developing Technical Reading Skills
M–3 Assist Students in Improving Their Writing Skills
M–4 Assist Students in Improving Their Oral Communication Skills
M–5 Assist Students in Improving Their Math Skills
M–6 Assist Students in Improving Their Survival Skills

RELATED PUBLICATIONS

Student Guide to Using Performance-Based Teacher Education Materials
Resource Person Guide to Using Performance-Based Teacher Education Materials
Guide to the Implementation of Performance-Based Teacher Education
Performance-Based Teacher Education: The State of the Art, General Education and Vocational Education

For information regarding availability and prices of these materials contact—AAVIM, American Association for Vocational Instructional Materials, 120 Driftmier Engineering Center, University of Georgia, Athens, Georgia 30602, (404) 542-2586

ISBN 0-89606-143-4